STIC-Biotech/ChemLib

175678 MG

From: Vivlemore, Tracy
Sent: Thursday, January 05, 2006 9:20 AM
To: STIC-Biotech/ChemLib
Subject: Sequence search request, application 10/028415

Hello,

For application 10/028,415, please perform a standard search of SEQ ID NO: 20.

Tracy Vivlemore PhD
Remsen 2B-02, AU 1635
Mailbox: 2C-18
Tel: 571-272-2914

****
Searcher:_____
Searcher Phone:_____
Date Searcher Picked up:_____
Date completed:_____
Searcher Prep Time:_____
Online Time:_____

****
Type of Search
NA#_____ AA#:_____
S/L:_____ Oligomer:_____
Encode/Transl:_____
Structure #:_____Text:____
Inventor:_____ Litigation:___

****
Vendors and cost where applicable
STN:_____
DIALOG:_____
QUESTEL/ORBIT:_____
LEXIS/NEXIS:_____
SEQUENCE SYSTEM:_____
WWW/Internet:_____
Other (Specify):_____

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 8, 2006, 17:18:19 ; Search time 102 Seconds
(without alignments)
383.396 Million cell updates/sec

Title: US-10-028-415-20
Perfect score: 22
Sequence: 1 ctgcacaggagggttggaatac 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	22	100.0	1481	3	US-09-429-323-3	Sequence 3, Appli
c	2	22	100.0	3073	2	US-07-688-352C-31	Sequence 31, Appl
c	3	22	100.0	3073	2	US-08-474-379C-31	Sequence 31, Appl
c	4	22	100.0	3073	3	US-09-146-249A-31	Sequence 31, Appl
c	5	22	100.0	3073	3	US-08-206-188B-31	Sequence 31, Appl
c	6	22	100.0	3073	6	PCT-US91-02714-30	Sequence 30, Appl
c	7	20	90.9	317	3	US-09-404-879A-380	Sequence 380, App
c	8	20	90.9	317	3	US-09-667-857-380	Sequence 380, App
c	9	20	90.9	317	3	US-10-198-053-380	Sequence 380, App
c	10	20	90.9	317	3	US-09-827-271-380	Sequence 380, App
c	11	17.2	78.2	601	3	US-09-949-016-33785	Sequence 33785, A
c	12	17.2	78.2	601	3	US-09-949-016-150908	Sequence 150908,
	13	17.2	78.2	2022	3	US-09-023-655-416	Sequence 416, App
	14	17.2	78.2	2088	3	US-09-949-016-4237	Sequence 4237, Ap
	15	17.2	78.2	4221	3	US-09-949-016-809	Sequence 809, App
	16	17.2	78.2	21593	3	US-09-949-016-15979	Sequence 15979, A
	17	17.2	78.2	96878	3	US-09-949-016-12551	Sequence 12551, A
c	18	17.2	78.2	238815	3	US-09-949-016-16274	Sequence 16274, A
	19	17.2	78.2	312474	3	US-09-949-016-17434	Sequence 17434, A
	20	16.8	76.4	68580	3	US-09-949-016-15844	Sequence 15844, A
c	21	16.8	76.4	87562	3	US-09-949-016-13685	Sequence 13685, A
	22	16.4	74.5	601	3	US-09-949-016-35086	Sequence 35086, A
	23	16.4	74.5	601	3	US-09-949-016-147990	Sequence 147990,
	24	16.4	74.5	1863	3	US-09-614-221A-53	Sequence 53, Appl
c	25	16.4	74.5	1969	3	US-09-949-016-4155	Sequence 4155, Ap
c	26	16.4	74.5	1996	3	US-09-949-016-873	Sequence 873, App
c	27	16.4	74.5	2194	2	US-08-948-569A-9	Sequence 9, Appli
c	28	16.4	74.5	2194	2	US-09-188-469-9	Sequence 9, Appli
c	29	16.4	74.5	2194	3	US-09-397-238A-9	Sequence 9, Appli
c	30	16.4	74.5	2330	3	US-10-104-047-1891	Sequence 1891, Ap
c	31	16.4	74.5	58821	3	US-09-949-016-15897	Sequence 15897, A
c	32	16.4	74.5	58824	3	US-09-949-016-12615	Sequence 12615, A
c	33	16.2	73.6	510	3	US-09-691-538A-8	Sequence 8, Appli
c	34	16.2	73.6	3568	3	US-10-160-719A-1	Sequence 1, Appli
c	35	16.2	73.6	3568	3	US-10-209-059-41	Sequence 41, Appl
c	36	16.2	73.6	3776	3	US-09-900-237-3	Sequence 3, Appli
c	37	16.2	73.6	77626	3	US-09-949-016-12608	Sequence 12608, A
	38	16.2	73.6	136264	3	US-09-949-016-12756	Sequence 12756, A
	39	16.2	73.6	136265	3	US-09-949-016-13001	Sequence 13001, A
	40	16	72.7	858	3	US-09-322-478-13	Sequence 13, Appl
	41	16	72.7	858	3	US-09-586-106D-13	Sequence 13, Appl
	42	16	72.7	858	3	US-10-799-870-13	Sequence 13, Appl
	43	16	72.7	9829	3	US-09-322-478-19	Sequence 19, Appl
	44	16	72.7	9829	3	US-09-586-106D-19	Sequence 19, Appl
	45	16	72.7	9829	3	US-10-799-870-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-09-429-323-3/c
; Sequence 3, Application US/09429323A
; Patent No. 6140126
; Patent No. 6140126 6140123
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Lex M. Cowsert
; TITLE OF INVENTION: ANTISENSE MODULATION OF Y-BOX BINDING PROTEIN 1 EXPRESSION
; FILE REFERENCE: RTS-0092
; CURRENT APPLICATION NUMBER: US/09/429,323A
; CURRENT FILING DATE: 1999-10-26
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 3
; LENGTH: 1481
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (127)..(1080)
US-09-429-323-3

Query Match 100.0%; Score 22; DB 3; Length 1481;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCACAGGAGGGTTGGAATAC 22
||||||||||||||||||||||
Db 768 CTGCACAGGAGGGTTGGAATAC 747

RESULT 2
US-07-688-352C-31/c
; Sequence 31, Application US/07688352C
; Patent No. 5527896
; GENERAL INFORMATION:
; APPLICANT: Wigler, Michael H.
; APPLICANT: Colicelli, John J.
; TITLE OF INVENTION: Cloning by Complementation and Related
; TITLE OF INVENTION: Processes
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago

```
;       STATE:  Illinois
;       COUNTRY:  USA
;       ZIP:  60603
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/07/688,352C
;       FILING DATE:  19910419
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/511,715
;       FILING DATE:  20-APR-1990
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Borun, Michael F.
;       REGISTRATION NUMBER:  25447
;       REFERENCE/DOCKET NUMBER:  27805/30197
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (312) 346-5750
;       TELEFAX:  (312) 984-9740
;       TELEX:  25-3856
;   INFORMATION FOR SEQ ID NO:  31:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  3073 base pairs
;       TYPE:  NUCLEIC ACID
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  3..1111
US-07-688-352C-31

  Query Match             100.0%;  Score 22;  DB 2;  Length 3073;
  Best Local Similarity   100.0%;  Pred. No. 0.25;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        779 CTGCACAGGAGGGTTGGAATAC 758


RESULT 3
US-08-474-379C-31/c
; Sequence 31, Application US/08474379C
; Patent No. 5977305
;  GENERAL INFORMATION:
;    APPLICANT:  Wigler, Michael H.
;    APPLICANT:  Colicelli, John J.
;    TITLE OF INVENTION:  CLONING BY COMPLEMENTATION AND RELATED
;    TITLE OF INVENTION:  PROCESSES
;    NUMBER OF SEQUENCES:  88
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  233 South Wacker Drive/6300 Sears Tower
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/474,379C
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/511,715
;      FILING DATE:  20-APR-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/206,188
;      FILING DATE:  01-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/688,352
;      FILING DATE:  19-APR-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clough, David W.
;      REGISTRATION NUMBER:  36,107
;      REFERENCE/DOCKET NUMBER:  27866/32771
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;  INFORMATION FOR SEQ ID NO:  31:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3073 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  3..1109
US-08-474-379C-31

  Query Match             100.0%;  Score 22;  DB 2;  Length 3073;
  Best Local Similarity   100.0%;  Pred. No. 0.25;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        779 CTGCACAGGAGGGTTGGAATAC 758


RESULT 4
US-09-146-249A-31/c
; Sequence 31, Application US/09146249A
; Patent No. 6069240
;  GENERAL INFORMATION:
;    APPLICANT:  Wigler, Michael H.
;    APPLICANT:  Colicelli, John J.
;    TITLE OF INVENTION:  Cloning by Complementation and Related
;    TITLE OF INVENTION:  Processes
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/146,249A
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/511,715
;      FILING DATE:  20-APR-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clough, David W.
;      REGISTRATION NUMBER:  36,107
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312-474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  31:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3073 base pairs
```

```
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  3..1111
US-09-146-249A-31

  Query Match             100.0%;  Score 22;  DB 3;  Length 3073;
  Best Local Similarity   100.0%;  Pred. No. 0.25;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        779 CTGCACAGGAGGGTTGGAATAC 758


RESULT 5
US-08-206-188B-31/c
; Sequence 31, Application US/08206188B
; Patent No. 6100025
;  GENERAL INFORMATION:
;    APPLICANT:  Wigler, Michael H.
;    APPLICANT:  Colicelli, John J.
;    TITLE OF INVENTION:  Cloning by Complementation and Related
;    TITLE OF INVENTION:  Processes
;    NUMBER OF SEQUENCES:  84
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/206,188B
;      FILING DATE:  01-MAR-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/511,715
;      FILING DATE:  20-APR-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Clough, David W.
;      REGISTRATION NUMBER:  36107
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312-474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  31:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3073 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  3..1111
US-08-206-188B-31

  Query Match             100.0%;  Score 22;  DB 3;  Length 3073;
  Best Local Similarity   100.0%;  Pred. No. 0.25;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        779 CTGCACAGGAGGGTTGGAATAC 758


RESULT 6
PCT-US91-02714-30/c
; Sequence 30, Application PC/TUS9102714
;  GENERAL INFORMATION:
;    APPLICANT:  Wigler, Michael H.
;    APPLICANT:  Colicelli, John J.
;    TITLE OF INVENTION:  Cloning by Complementation and Related
;    TITLE OF INVENTION:  Processes
;    NUMBER OF SEQUENCES:  55
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray &
;      ADDRESSEE:  Bicknell
;      STREET:  Two First National Plaza, 20 South Clark
;      STREET:  Street
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60603
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US91/02714
;      FILING DATE:  19910419
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/511,715
;      FILING DATE:  20-APR-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Borun, Michael F.
;      REGISTRATION NUMBER:  25447
;      REFERENCE/DOCKET NUMBER:  27805/30197
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 346-5750
;      TELEFAX:  (312) 984-9740
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  30:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3073 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  3..1111
PCT-US91-02714-30

  Query Match             100.0%;  Score 22;  DB 6;  Length 3073;
  Best Local Similarity   100.0%;  Pred. No. 0.25;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        779 CTGCACAGGAGGGTTGGAATAC 758


RESULT 7
US-09-404-879A-380/c
; Sequence 380, Application US/09404879A
; Patent No. 6468546
; GENERAL INFORMATION:
;  APPLICANT:  Mitcham, Jennifer L.
;  APPLICANT:  King, Gordon E.
;  APPLICANT:  Algate, Paul A.
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE:  210121.462C2
;  CURRENT APPLICATION NUMBER: US/09/404,879A
```

```
;  CURRENT FILING DATE:  1999-09-24
;  NUMBER OF SEQ ID NOS:  393
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 380
;   LENGTH: 317
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(317)
;   OTHER INFORMATION: n = A,T,C or G
US-09-404-879A-380

  Query Match             90.9%;  Score 20;  DB 3;  Length 317;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 GCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||
Db         29 GCACAGGAGGGTTGGAATAC 10


RESULT 8
US-09-667-857-380/c
; Sequence 380, Application US/09667857
; Patent No. 6699664
; GENERAL INFORMATION:
;  APPLICANT: Mitcham, Jennifer L.
;  APPLICANT:  King, Gordon E.
;  APPLICANT:  Algate, Paul A.
;  APPLICANT:  Fling, Steven P.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Fanger, Gary Richard
;  APPLICANT:  Reed, Steven G.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C5
;  CURRENT APPLICATION NUMBER: US/09/667,857
;  CURRENT FILING DATE:  2000-09-20
;  NUMBER OF SEQ ID NOS: 455
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 380
;   LENGTH: 317
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(317)
;   OTHER INFORMATION: n = A,T,C or G
US-09-667-857-380

  Query Match             90.9%;  Score 20;  DB 3;  Length 317;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 GCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||
Db         29 GCACAGGAGGGTTGGAATAC 10


RESULT 9
US-10-198-053-380/c
; Sequence 380, Application US/10198053
; Patent No. 6858710
; GENERAL INFORMATION:
;  APPLICANT: Bangur, Chaitanya S.
;  APPLICANT:  Retter, Marc W.
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Hill, Paul
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C9
;  CURRENT APPLICATION NUMBER: US/10/198,053
;  CURRENT FILING DATE:  2002-07-17
;  NUMBER OF SEQ ID NOS: 624
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 380
;   LENGTH: 317
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 30, 32
;   OTHER INFORMATION: n = A,T,C or G
US-10-198-053-380

  Query Match             90.9%;  Score 20;  DB 3;  Length 317;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 GCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||
Db         29 GCACAGGAGGGTTGGAATAC 10


RESULT 10
US-09-827-271-380/c
; Sequence 380, Application US/09827271
; Patent No. 6962980
; GENERAL INFORMATION:
;  APPLICANT: Retter, Marc W.
;  APPLICANT:  Fanger, Gary R.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF OVARIAN CANCER
;  FILE REFERENCE: 210121.462C6
;  CURRENT APPLICATION NUMBER: US/09/827,271
;  CURRENT FILING DATE:  2001-04-04
;  NUMBER OF SEQ ID NOS: 461
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 380
;   LENGTH: 317
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(317)
;   OTHER INFORMATION: n = A,T,C or G
US-09-827-271-380

  Query Match             90.9%;  Score 20;  DB 3;  Length 317;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          3 GCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||
Db         29 GCACAGGAGGGTTGGAATAC 10


RESULT 11
US-09-949-016-33785/c
; Sequence 33785, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
```

```
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 33785
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-33785

  Query Match             78.2%;  Score 17.2;  DB 3;  Length 601;
  Best Local Similarity   86.4%;  Pred. No. 48;
  Matches   19;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              |||| |||| ||||||||| ||
Db        409 CTGCCCAGGTGGGTTGGAACAC 388


RESULT 12
US-09-949-016-150908/c
; Sequence 150908, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 150908
;   LENGTH: 601
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-150908

  Query Match             78.2%;  Score 17.2;  DB 3;  Length 601;
  Best Local Similarity   86.4%;  Pred. No. 48;
  Matches   19;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              |||| |||| ||||||||| ||
Db        409 CTGCCCAGGTGGGTTGGAACAC 388


RESULT 13
US-09-023-655-416
; Sequence 416, Application US/09023655
; Patent No. 6607879
;  GENERAL INFORMATION:
;    APPLICANT:  Cocks, Benjamin G.
;    APPLICANT:  Susan G. Stuart
;    APPLICANT:  Jeffrey J. Seilhamer
;    TITLE OF INVENTION:  COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
;    TITLE OF INVENTION:  EXPRESSION
;    NUMBER OF SEQUENCES:  1508
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  INCYTE PHARMACEUTICALS, INC.
;      STREET:  3174 PORTER DRIVE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Word Perfect 6.1 for Windows/MS-DOS 6.2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/023,655
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Zeller, Karen J.
;      REGISTRATION NUMBER:  37,071
;      REFERENCE/DOCKET NUMBER:  PA-0001 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 855-0555
;      TELEFAX:  (650) 845-4166
;  INFORMATION FOR SEQ ID NO:  416:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2022 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  BLADNOT03
;      CLONE:  1602090
US-09-023-655-416

  Query Match             78.2%;  Score 17.2;  DB 3;  Length 2022;
  Best Local Similarity   86.4%;  Pred. No. 59;
  Matches   19;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              |||| |||| ||||||||| ||
Db       1223 CTGCCCAGGTGGGTTGGAACAC 1244


RESULT 14
US-09-949-016-4237
; Sequence 4237, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4237
;   LENGTH: 2088
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-4237

  Query Match             78.2%;  Score 17.2;  DB 3;  Length 2088;
  Best Local Similarity   86.4%;  Pred. No. 59;
  Matches   19;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              |||| |||| ||||||||| ||
Db       1558 CTGCCCAGGTGGGTTGGAACAC 1579


RESULT 15
US-09-949-016-809
```

```
; Sequence 809, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 809
;   LENGTH: 4221
;   TYPE: DNA
;   ORGANISM: Human
US-09-949-016-809

  Query Match             78.2%;  Score 17.2;  DB 3;  Length 4221;
  Best Local Similarity   86.4%;  Pred. No. 67;
  Matches   19;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              |||| |||| ||||||||| ||
Db       3690 CTGCCCAGGTGGGTTGGAACAC 3711


Search completed: January  8, 2006, 18:40:23
Job time : 105 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 8, 2006, 17:22:30 ; Search time 561 Seconds
(without alignments)
324.289 Million cell updates/sec

Title: US-10-028-415-20
Perfect score: 22
Sequence: 1 ctgcacaggagggttggaatac 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	22	100.0	22	5	US-10-028-415-20	Sequence 20, Appl
c	2	22	100.0	254	8	US-10-425-115-179020	Sequence 179020,
	3	22	100.0	314	6	US-10-029-386-17873	Sequence 17873, A
	4	22	100.0	400	5	US-10-040-739-1	Sequence 1, Appli
	5	22	100.0	451	3	US-09-969-034-3986	Sequence 3986, Ap
c	6	22	100.0	478	3	US-09-918-995-36678	Sequence 36678, A
	7	22	100.0	519	5	US-10-027-632-16296	Sequence 16296, A
	8	22	100.0	519	5	US-10-027-632-16297	Sequence 16297, A
	9	22	100.0	519	5	US-10-027-632-16298	Sequence 16298, A
	10	22	100.0	519	6	US-10-027-632-16296	Sequence 16296, A
	11	22	100.0	519	6	US-10-027-632-16297	Sequence 16297, A
	12	22	100.0	519	6	US-10-027-632-16298	Sequence 16298, A
c	13	22	100.0	560	3	US-09-969-034-1551	Sequence 1551, Ap
	14	22	100.0	595	6	US-10-029-386-4173	Sequence 4173, Ap
	15	22	100.0	822	6	US-10-257-826A-285	Sequence 285, App
c	16	22	100.0	954	3	US-09-972-331-26	Sequence 26, Appl
c	17	22	100.0	1474	6	US-10-172-118-1057	Sequence 1057, Ap
c	18	22	100.0	1474	6	US-10-388-360-332	Sequence 332, App
c	19	22	100.0	1474	7	US-10-342-887-1057	Sequence 1057, Ap
c	20	22	100.0	1474	9	US-10-956-157-1265	Sequence 1265, Ap
c	21	22	100.0	1481	3	US-09-880-107-2100	Sequence 2100, Ap
c	22	22	100.0	1481	5	US-10-023-969-2	Sequence 2, Appli
c	23	22	100.0	1481	6	US-10-388-215-3	Sequence 3, Appli
c	24	22	100.0	1516	9	US-10-483-505-45	Sequence 45, Appl
c	25	22	100.0	1520	3	US-09-971-392-82	Sequence 82, Appl
c	26	22	100.0	1554	9	US-10-887-553A-366	Sequence 366, App
c	27	20.4	92.7	448	3	US-09-960-352-10627	Sequence 10627, A
c	28	20.4	92.7	1400	9	US-10-956-157-9444	Sequence 9444, Ap
c	29	20.4	92.7	1522	9	US-10-956-157-4209	Sequence 4209, Ap
c	30	20	90.9	317	3	US-09-884-441-380	Sequence 380, App
c	31	20	90.9	317	3	US-09-907-969-380	Sequence 380, App
c	32	20	90.9	317	3	US-09-827-271-380	Sequence 380, App
c	33	20	90.9	317	6	US-10-198-053-380	Sequence 380, App
c	34	20	90.9	317	8	US-10-860-790-380	Sequence 380, App
	35	19.4	88.2	642	7	US-10-767-701-5414	Sequence 5414, Ap
c	36	19.4	88.2	1189	8	US-10-425-115-36622	Sequence 36622, A
c	37	19.4	88.2	1489	6	US-10-388-215-52	Sequence 52, Appl
c	38	19.4	88.2	1489	7	US-10-152-319A-1916	Sequence 1916, Ap
c	39	19	86.4	2102	9	US-10-450-763-18807	Sequence 18807, A
c	40	18.8	85.5	368	3	US-09-960-352-10869	Sequence 10869, A
	41	17.4	79.1	576	4	US-09-925-065A-146954	Sequence 146954,
c	42	17.4	79.1	192992	9	US-10-461-862-95	Sequence 95, Appl
	43	17.2	78.2	557	4	US-09-925-065A-158278	Sequence 158278,
c	44	17.2	78.2	578	9	US-10-779-543-20742	Sequence 20742, A
c	45	17.2	78.2	596	4	US-09-925-065A-524409	Sequence 524409,

ALIGNMENTS

RESULT 1
US-10-028-415-20
; Sequence 20, Application US/10028415
; Publication No. US20020151063A1
; GENERAL INFORMATION:
; APPLICANT: Lasham, Annette
; APPLICANT: Watson, James D.
; TITLE OF INVENTION: Methods for Modulating Apoptotic Cell
; TITLE OF INVENTION: Death
; FILE REFERENCE: 11000.1004c3
; CURRENT APPLICATION NUMBER: US/10/028,415
; CURRENT FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: PCT/NZ01/00286
; PRIOR FILING DATE: 2001-11-28
; PRIOR APPLICATION NUMBER: US 09/724,809
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 09/036,004
; PRIOR FILING DATE: 1998-03-04
; PRIOR APPLICATION NUMBER: US 08/713,557
; PRIOR FILING DATE: 1996-08-30
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 22
; TYPE: DNA
; ORGANISM: Human
US-10-028-415-20

Query Match 100.0%; Score 22; DB 5; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCACAGGAGGGTTGGAATAC 22
||||||||||||||||||||||
Db 1 CTGCACAGGAGGGTTGGAATAC 22

RESULT 2
US-10-425-115-179020/c
; Sequence 179020, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei

```
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(53222)B
;  CURRENT APPLICATION NUMBER: US/10/425,115
;  CURRENT FILING DATE:  2003-04-28
;  NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 179020
;   LENGTH: 254
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_9484C.1
US-10-425-115-179020

  Query Match             100.0%;  Score 22;  DB 8;  Length 254;
  Best Local Similarity   100.0%;  Pred. No. 0.52;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        154 CTGCACAGGAGGGTTGGAATAC 133


RESULT 3
US-10-029-386-17873
; Sequence 17873, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
;  TITLE OF INVENTION:  EXPRESSION ANALYSIS TWO
;  FILE REFERENCE: AEOMICA-X-2
;  CURRENT APPLICATION NUMBER: US/10/029,386
;  CURRENT FILING DATE:  2001-12-20
;  NUMBER OF SEQ ID NOS: 34288
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 17873
;   LENGTH: 314
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL138788.1
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL =  12
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL =  12
;   OTHER INFORMATION: NT HIT: gi14742436, EVALUE 0.00e+00
;   OTHER INFORMATION: SWISSPROT HIT: P27817, EVALUE 3.00e-09
;   OTHER INFORMATION: EST_HUMAN HIT: BG676581.1, EVALUE 0.00e+00
US-10-029-386-17873

  Query Match             100.0%;  Score 22;  DB 6;  Length 314;
  Best Local Similarity   100.0%;  Pred. No. 0.52;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db         43 CTGCACAGGAGGGTTGGAATAC 64


RESULT 4
US-10-040-739-1
; Sequence 1, Application US/10040739
; Publication No. US20020173635A1
;   GENERAL INFORMATION:
;        APPLICANT: Jacobs, Kenneth
;                   McCoy, John
;                   LaVallie, Edward
;                   Racie, Lisa
;                   Merberg, David
;                   Treacy, Maurice
;                   Spaulding, Vikki
;        TITLE OF INVENTION: SECRETED, EXPRESSED SEQUENCE TAGS
;        NUMBER OF SEQUENCES: 1519
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Genetics Institute, Inc.
;             STREET: 87 CambridgePark Drive
;             CITY: Cambridge
;             STATE: Massachusetts
;             COUNTRY: U.S.A
;             ZIP: 02140
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy Disk
;             COMPUTER: IBM PC Compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/040,739
;             FILING DATE: 07-Jan-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/036,520
;             FILING DATE: 03-JUN-1998
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Brown, Scott A.
;             REGISTRATION NUMBER: 32,724
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617) 498-8224
;             TELEFAX: (617) 876-5851
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 400 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        MOLECULE TYPE: cDNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-040-739-1

  Query Match             100.0%;  Score 22;  DB 5;  Length 400;
  Best Local Similarity   100.0%;  Pred. No. 0.52;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        200 CTGCACAGGAGGGTTGGAATAC 221


RESULT 5
US-09-969-034-3986
; Sequence 3986, Application US/09969034
; Publication No. US20040110668A1
; GENERAL INFORMATION:
;  APPLICANT: Burgess, Christopher C.
;  APPLICANT:  Astle, Jon H.
;  APPLICANT:  Carroll, Eddie III
;  APPLICANT:  Catino, Theodore J.
;  APPLICANT:  Dwivedi, Poornima
;  APPLICANT:  Molino, Gary A.
;  APPLICANT:  Thiagalingam, Arunthathi
;  APPLICANT:  Lewis, Marcia E.
;  TITLE OF INVENTION: Nucleic Acid Sequences Differentially
;  TITLE OF INVENTION:  Expressed in Cancer Tissue
;  FILE REFERENCE: 1657/1032
;  CURRENT APPLICATION NUMBER: US/09/969,034
;  CURRENT FILING DATE:  2001-10-02
;  PRIOR APPLICATION NUMBER: 60/237,271
;  PRIOR FILING DATE: 2000-02-10
;  NUMBER OF SEQ ID NOS: 4494
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3986
;   LENGTH: 451
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
```

```
;   LOCATION: 418, 420
;   OTHER INFORMATION: n = A,T,C or G
US-09-969-034-3986

  Query Match             100.0%;  Score 22;  DB 3;  Length 451;
  Best Local Similarity   100.0%;  Pred. No. 0.53;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        200 CTGCACAGGAGGGTTGGAATAC 221


RESULT 6
US-09-918-995-36678/c
; Sequence 36678, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 36678
;   LENGTH: 478
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(478)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-36678

  Query Match             100.0%;  Score 22;  DB 3;  Length 478;
  Best Local Similarity   100.0%;  Pred. No. 0.53;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        101 CTGCACAGGAGGGTTGGAATAC 80


RESULT 7
US-10-027-632-16296
; Sequence 16296, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16296
;   LENGTH: 519
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-16296

  Query Match             100.0%;  Score 22;  DB 5;  Length 519;
  Best Local Similarity   100.0%;  Pred. No. 0.53;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        226 CTGCACAGGAGGGTTGGAATAC 247


RESULT 8
US-10-027-632-16297
; Sequence 16297, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16297
;   LENGTH: 519
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-16297

  Query Match             100.0%;  Score 22;  DB 5;  Length 519;
  Best Local Similarity   100.0%;  Pred. No. 0.53;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        226 CTGCACAGGAGGGTTGGAATAC 247


RESULT 9
US-10-027-632-16298
; Sequence 16298, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
```

```
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16298
;   LENGTH: 519
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-16298

  Query Match             100.0%;  Score 22;  DB 5;  Length 519;
  Best Local Similarity   100.0%;  Pred. No. 0.53;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        226 CTGCACAGGAGGGTTGGAATAC 247


RESULT 10
US-10-027-632-16296
; Sequence 16296, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16296
;   LENGTH: 519
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-16296

  Query Match             100.0%;  Score 22;  DB 6;  Length 519;
  Best Local Similarity   100.0%;  Pred. No. 0.53;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        226 CTGCACAGGAGGGTTGGAATAC 247


RESULT 11
US-10-027-632-16297
; Sequence 16297, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16297
;   LENGTH: 519
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-16297

  Query Match             100.0%;  Score 22;  DB 6;  Length 519;
  Best Local Similarity   100.0%;  Pred. No. 0.53;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        226 CTGCACAGGAGGGTTGGAATAC 247


RESULT 12
US-10-027-632-16298
; Sequence 16298, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
;  APPLICANT: Wang, David G.
;  TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
;  TITLE OF INVENTION:  Polymorphisms in the Human Genome
;  FILE REFERENCE: 108827.129
;  CURRENT APPLICATION NUMBER: US/10/027,632
;  CURRENT FILING DATE:  2002-04-30
;  PRIOR APPLICATION NUMBER: US 60/218,006
;  PRIOR FILING DATE: 2000-07-12
;  PRIOR APPLICATION NUMBER: US 60/198,676
;  PRIOR FILING DATE: 2000-04-20
;  PRIOR APPLICATION NUMBER: US 60/193,483
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: US 60/185,218
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: US 60/167,363
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: US 60/156,358
;  PRIOR FILING DATE: 1999-09-28
;  PRIOR APPLICATION NUMBER: US 60/146,002
;  PRIOR FILING DATE: 1999-08-09
;  NUMBER OF SEQ ID NOS: 325720
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16298
;   LENGTH: 519
;   TYPE: DNA
;   ORGANISM: Human
US-10-027-632-16298

  Query Match             100.0%;  Score 22;  DB 6;  Length 519;
  Best Local Similarity   100.0%;  Pred. No. 0.53;
```

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        226 CTGCACAGGAGGGTTGGAATAC 247
```

RESULT 13
US-09-969-034-1551/c
; Sequence 1551, Application US/09969034
; Publication No. US20040110668A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Astle, Jon H.
; APPLICANT: Carroll, Eddie III
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dwivedi, Poornima
; APPLICANT: Molino, Gary A.
; APPLICANT: Thiagalingam, Arunthathi
; APPLICANT: Lewis, Marcia E.
; TITLE OF INVENTION: Nucleic Acid Sequences Differentially
; TITLE OF INVENTION: Expressed in Cancer Tissue
; FILE REFERENCE: 1657/1032
; CURRENT APPLICATION NUMBER: US/09/969,034
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/237,271
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 4494
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1551
; LENGTH: 560
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 385, 394, 473, 528, 560
; OTHER INFORMATION: n = A,T,C or G
US-09-969-034-1551

Query Match 100.0%; Score 22; DB 3; Length 560;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        348 CTGCACAGGAGGGTTGGAATAC 327
```

RESULT 14
US-10-029-386-4173
; Sequence 4173, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 4173
; LENGTH: 595
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL138788.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 12
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 12
; OTHER INFORMATION: SWISSPROT HIT: P27817, EVALUE 2.00e-12
; OTHER INFORMATION: NT HIT: gi14742436, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: AI937888.1, EVALUE 0.00e+00
US-10-029-386-4173

Query Match 100.0%; Score 22; DB 6; Length 595;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        299 CTGCACAGGAGGGTTGGAATAC 320
```

RESULT 15
US-10-257-826A-285
; Sequence 285, Application US/10257826A
; Publication No. US20030181407A1
; GENERAL INFORMATION:
; APPLICANT: SA MAJESTE LA REINE DU CHEF DU CANADA
; APPLICANT: PALIN, Marie-France
; APPLICANT: POMAR, Candido
; APPLICANT: GARIEPY, Claude
; TITLE OF INVENTION: Steatosis-modulating factors and uses
; TITLE OF INVENTION: thereof
; FILE REFERENCE: 14654-2US
; CURRENT APPLICATION NUMBER: US/10/257,826A
; CURRENT FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: 60/197936
; PRIOR FILING DATE: 2000-04-17
; PRIOR APPLICATION NUMBER: PCT/CA01/00509
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 285
; LENGTH: 822
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial sequence
; OTHER INFORMATION: Muscular steatosis
; OTHER INFORMATION: Porcine
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(822)
; OTHER INFORMATION: n = A,T,C or G
US-10-257-826A-285

Query Match 100.0%; Score 22; DB 6; Length 822;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db         71 CTGCACAGGAGGGTTGGAATAC 92
```

Search completed: January 8, 2006, 18:49:55
Job time : 562 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 8, 2006, 17:27:39 ; Search time 991 Seconds
(without alignments)
16.185 Million cell updates/sec

Title: US-10-028-415-20
Perfect score: 22
Sequence: 1 ctgcacaggagggttggaatac 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4637633 seqs, 364532575 residues

Total number of hits satisfying chosen parameters: 9275266

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA_New:*
1: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq2:*
9: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq3:*
10: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	22	100.0	25	7	US-11-121-849-531868	Sequence 531868,
	2	20.4	92.7	1360	6	US-10-750-185-47669	Sequence 47669, A
	3	20.4	92.7	1360	6	US-10-750-623-47669	Sequence 47669, A
c	4	19	86.4	25	7	US-11-121-849-531869	Sequence 531869,
c	5	19	86.4	1400	7	US-11-136-527-4432	Sequence 4432, Ap
c	6	19	86.4	1666	7	US-11-136-527-336	Sequence 336, App
	7	17.2	78.2	1882	6	US-10-750-185-40265	Sequence 40265, A
	8	17.2	78.2	1882	6	US-10-750-623-40265	Sequence 40265, A
	9	16.8	76.4	75330	7	US-11-124-368A-2915	Sequence 2915, Ap
c	10	16.4	74.5	201	7	US-11-124-368A-6677	Sequence 6677, Ap
	11	16.4	74.5	201	7	US-11-124-368A-15319	Sequence 15319, A
	12	16.4	74.5	100000	7	US-11-124-368A-2885	Sequence 2885, Ap
	13	16.2	73.6	25	7	US-11-069-908-552	Sequence 552, App
	14	16.2	73.6	25	7	US-11-069-908-2918	Sequence 2918, Ap
	15	15.8	71.8	3154	7	US-11-108-528-45	Sequence 45, Appl
	16	15.8	71.8	3355	6	US-10-750-185-61156	Sequence 61156, A
	17	15.8	71.8	3355	6	US-10-750-623-61156	Sequence 61156, A
c	18	15.8	71.8	42823	7	US-11-066-725-18	Sequence 18, Appl
	19	15.8	71.8	56054	6	US-10-995-561-13402	Sequence 13402, A
	20	15.6	70.9	201	6	US-10-995-561-80639	Sequence 80639, A
	21	15.6	70.9	201	6	US-10-995-561-80761	Sequence 80761, A
	22	15.6	70.9	201	6	US-10-995-561-80786	Sequence 80786, A
	23	15.6	70.9	745	6	US-10-750-185-60160	Sequence 60160, A
	24	15.6	70.9	745	6	US-10-750-623-60160	Sequence 60160, A
c	25	15.6	70.9	823	6	US-10-750-185-61921	Sequence 61921, A
c	26	15.6	70.9	823	6	US-10-750-623-61921	Sequence 61921, A
c	27	15.6	70.9	1020	6	US-10-750-185-53026	Sequence 53026, A
c	28	15.6	70.9	1020	6	US-10-750-623-53026	Sequence 53026, A
c	29	15.6	70.9	1101	6	US-10-750-185-49755	Sequence 49755, A
c	30	15.6	70.9	1101	6	US-10-750-623-49755	Sequence 49755, A
	31	15.6	70.9	1149	6	US-10-750-185-59099	Sequence 59099, A
	32	15.6	70.9	1149	6	US-10-750-623-59099	Sequence 59099, A
	33	15.6	70.9	1274	6	US-10-750-185-25555	Sequence 25555, A
	34	15.6	70.9	1274	6	US-10-750-623-25555	Sequence 25555, A
	35	15.6	70.9	1298	6	US-10-750-185-27228	Sequence 27228, A
	36	15.6	70.9	1298	6	US-10-750-623-27228	Sequence 27228, A
	37	15.6	70.9	65885	6	US-10-995-561-13490	Sequence 13490, A
	38	15.6	70.9	1082144	7	US-11-117-187-211	Sequence 211, App
c	39	15.4	70.0	1631	7	US-11-108-528-55	Sequence 55, Appl
c	40	15.4	70.0	2586	6	US-10-750-185-24808	Sequence 24808, A
c	41	15.4	70.0	2586	6	US-10-750-623-24808	Sequence 24808, A
	42	15.4	70.0	100000	7	US-11-124-368A-2906	Sequence 2906, Ap
c	43	15.4	70.0	138821	7	US-11-121-086-80	Sequence 80, Appl
c	44	15.4	70.0	155515	7	US-11-112-908-42	Sequence 42, Appl
c	45	15.4	70.0	159660	7	US-11-112-908-43	Sequence 43, Appl

ALIGNMENTS

RESULT 1
US-11-121-849-531868/c
; Sequence 531868, Application US/11121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:
; APPLICANT: John Palma
; TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded Sa
; TITLE OF INVENTION: Microarrays
; FILE REFERENCE: 3684.1
; CURRENT APPLICATION NUMBER: US/11/121,849
; CURRENT FILING DATE: 2005-05-03
; PRIOR APPLICATION NUMBER: 60/567,949
; PRIOR FILING DATE: 2004-05-03
; NUMBER OF SEQ ID NOS: 673904
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 531868
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-11-121-849-531868

Query Match 100.0%; Score 22; DB 7; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.041;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db         23 CTGCACAGGAGGGTTGGAATAC 2
```

RESULT 2
US-10-750-185-47669
; Sequence 47669, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMI1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31

; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 47669
; LENGTH: 1360
; TYPE: DNA
; ORGANISM: Bovine 19866880926144
US-10-750-185-47669

Query Match 92.7%; Score 20.4; DB 6; Length 1360;
Best Local Similarity 95.5%; Pred. No. 0.36;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTGCACAGGAGGGTTGGAATAC 22
||||||||||||||||||| |||
Db 242 CTGCACAGGAGGGTTGGAGTAC 263

RESULT 3
US-10-750-623-47669
; Sequence 47669, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMI1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 47669
; LENGTH: 1360
; TYPE: DNA
; ORGANISM: Bovine 19866880926144
US-10-750-623-47669

Query Match 92.7%; Score 20.4; DB 6; Length 1360;
Best Local Similarity 95.5%; Pred. No. 0.36;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CTGCACAGGAGGGTTGGAATAC 22
||||||||||||||||||| |||
Db 242 CTGCACAGGAGGGTTGGAGTAC 263

RESULT 4
US-11-121-849-531869/c
; Sequence 531869, Application US/11121849
; Publication No. US20050272080A1
; GENERAL INFORMATION:
; APPLICANT: John Palma
; TITLE OF INVENTION: Methods of Genetic Analysis of Formalin Fixed Paraffin Embedded S
; TITLE OF INVENTION: Microarrays
; FILE REFERENCE: 3684.1
; CURRENT APPLICATION NUMBER: US/11/121,849
; CURRENT FILING DATE: 2005-05-03
; PRIOR APPLICATION NUMBER: 60/567,949
; PRIOR FILING DATE: 2004-05-03
; NUMBER OF SEQ ID NOS: 673904
; SOFTWARE: Microarray Probe Sequence Listing Generator V 1.1
; SEQ ID NO 531869
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Homo sapien
US-11-121-849-531869

Query Match 86.4%; Score 19; DB 7; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCACAGGAGGGTTGGAA 19
|||||||||||||||||||
Db 19 CTGCACAGGAGGGTTGGAA 1

RESULT 5
US-11-136-527-4432/c
; Sequence 4432, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4432
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1375)..(1375)
; OTHER INFORMATION: n is a, c, g, or t
US-11-136-527-4432

Query Match 86.4%; Score 19; DB 7; Length 1400;
Best Local Similarity 90.5%; Pred. No. 1.8;
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TGCACAGGAGGGTTGGAATAC 22
|||||||| |||||||:||||
Db 630 TGCACAGGGGGGTTGGRATAC 610

RESULT 6
US-11-136-527-336/c
; Sequence 336, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 336
; LENGTH: 1666
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1641)..(1641)
; OTHER INFORMATION: n is a, c, g, or t
US-11-136-527-336

Query Match 86.4%; Score 19; DB 7; Length 1666;
Best Local Similarity 90.5%; Pred. No. 1.9;
Matches 19; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy          2 TGCACAGGAGGGTTGGAATAC 22
              |||||||| ||||||| ||||
Db        896 TGCACAGGGGGGTTGGRATAC 876
```

RESULT 7
US-10-750-185-40265
; Sequence 40265, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMI1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 40265
; LENGTH: 1882
; TYPE: DNA
; ORGANISM: Bovine 19866880841243
US-10-750-185-40265

Query Match 78.2%; Score 17.2; DB 6; Length 1882;
Best Local Similarity 86.4%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              | ||||||||||  ||||||||
Db        764 CAGCACAGGAGGAATGGAATAC 785
```

RESULT 8
US-10-750-623-40265
; Sequence 40265, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMI1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 40265
; LENGTH: 1882
; TYPE: DNA
; ORGANISM: Bovine 19866880841243
US-10-750-623-40265

Query Match 78.2%; Score 17.2; DB 6; Length 1882;
Best Local Similarity 86.4%; Pred. No. 15;
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              | ||||||||||  ||||||||
Db        764 CAGCACAGGAGGAATGGAATAC 785
```

RESULT 9
US-11-124-368A-2915
; Sequence 2915, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2915
; LENGTH: 75330
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-124-368A-2915

Query Match 76.4%; Score 16.8; DB 7; Length 75330;
Best Local Similarity 90.0%; Pred. No. 33;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          3 GCACAGGAGGGTTGGAATAC 22
              |||||||||| ||||||| |
Db      18771 GCACAGGAGGCTTGGAATCC 18790
```

RESULT 10
US-11-124-368A-6677/c
; Sequence 6677, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6677
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-124-368A-6677

Query Match 74.5%; Score 16.4; DB 7; Length 201;
Best Local Similarity 94.4%; Pred. No. 33;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          2 TGCACAGGAGGGTTGGAA 19
              |||||||||| ||||||
Db         37 TGCACAGGAGGTTTGGAA 20
```

RESULT 11
US-11-124-368A-15319

; Sequence 15319, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15319
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-124-368A-15319

Query Match 74.5%; Score 16.4; DB 7; Length 201;
Best Local Similarity 94.4%; Pred. No. 33;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GCACAGGAGGGTTGGAAT 20
||||||||||| |||||||
Db 184 GCACAGGAGGCTTGGAAT 201

RESULT 12
US-11-124-368A-2885
; Sequence 2885, Application US/11124368A
; Publication No. US20050287559A1
; GENERAL INFORMATION:
; APPLICANT: Michele Cargill
; APPLICANT: James J. Devlin
; APPLICANT: May Luke
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Vascular Diseases, Methods of Detection and Uses Thereof
; FILE REFERENCE: CL001524
; CURRENT APPLICATION NUMBER: US/11/124,368A
; CURRENT FILING DATE: 2005-05-09
; PRIOR APPLICATION NUMBER: US 60/568,845
; PRIOR FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: US 60/625,936
; PRIOR FILING DATE: 2004-11-09
; NUMBER OF SEQ ID NOS: 21112
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2885
; LENGTH: 100000
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-124-368A-2885

Query Match 74.5%; Score 16.4; DB 7; Length 100000;
Best Local Similarity 94.4%; Pred. No. 53;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 TGCACAGGAGGGTTGGAA 19
||||||||||| ||||||
Db 25636 TGCACAGGAGGTTTGGAA 25653

RESULT 13
US-11-069-908-552
; Sequence 552, Application US/11069908
; Publication No. US20050266432A1
; GENERAL INFORMATION:
; APPLICANT: OLIPHANT, ARNOLD
; APPLICANT: MURRAY, SARAH
; TITLE OF INVENTION: HAPLOTYPE MARKERS FOR DIAGNOSING SUSCEPTIBILITY TO IMMUNOLOGICAL
; TITLE OF INVENTION: CONDITIONS
; FILE REFERENCE: 029011-0402
; CURRENT APPLICATION NUMBER: US/11/069,908
; CURRENT FILING DATE: 2005-02-28
; PRIOR APPLICATION NUMBER: 60/547,823
; PRIOR FILING DATE: 2004-02-26
; NUMBER OF SEQ ID NOS: 7098
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 552
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide
US-11-069-908-552

Query Match 73.6%; Score 16.2; DB 7; Length 25;
Best Local Similarity 85.7%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 TGCACAGGAGGGTTGGAATAC 22
|| ||||||| |||||||||
Db 1 TGTACAGGAGCTTTGGAATAC 21

RESULT 14
US-11-069-908-2918
; Sequence 2918, Application US/11069908
; Publication No. US20050266432A1
; GENERAL INFORMATION:
; APPLICANT: OLIPHANT, ARNOLD
; APPLICANT: MURRAY, SARAH
; TITLE OF INVENTION: HAPLOTYPE MARKERS FOR DIAGNOSING SUSCEPTIBILITY TO IMMUNOLOGICAL
; TITLE OF INVENTION: CONDITIONS
; FILE REFERENCE: 029011-0402
; CURRENT APPLICATION NUMBER: US/11/069,908
; CURRENT FILING DATE: 2005-02-28
; PRIOR APPLICATION NUMBER: 60/547,823
; PRIOR FILING DATE: 2004-02-26
; NUMBER OF SEQ ID NOS: 7098
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2918
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide
US-11-069-908-2918

Query Match 73.6%; Score 16.2; DB 7; Length 25;
Best Local Similarity 85.7%; Pred. No. 36;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 TGCACAGGAGGGTTGGAATAC 22
|| ||||||| |||||||||
Db 1 TGTACAGGAGCTTTGGAATAC 21

RESULT 15
US-11-108-528-45
; Sequence 45, Application US/11108528
; Publication No. US20050261189A1
; GENERAL INFORMATION:
; APPLICANT: Larsen, Glenn
; APPLICANT: Marvin, Martha
; APPLICANT: Li, Dean Y.
; APPLICANT: Wang, Elizabeth
; APPLICANT: Chen, C. M. Amy
; APPLICANT: Shamah, Steven M.
; TITLE OF INVENTION: METHODS OF PROMOTING CARDIAC CELL
; TITLE OF INVENTION: PROLIFERATION
; FILE REFERENCE: HYDR-P01-041

; CURRENT APPLICATION NUMBER: US/11/108,528
; CURRENT FILING DATE: 2005-04-18
; PRIOR APPLICATION NUMBER: US 60/563,137
; PRIOR FILING DATE: 2004-04-16
; PRIOR APPLICATION NUMBER: US 60/598,368
; PRIOR FILING DATE: 2004-08-02
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 3154
; TYPE: DNA
; ORGANISM: Mouse
US-11-108-528-45

Query Match 71.8%; Score 15.8; DB 7; Length 3154;
Best Local Similarity 89.5%; Pred. No. 82;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          3 GCACAGGAGGGTTGGAATA 21
              ||||||||||| || |||
Db       1277 GCACAGGAGGGTGGGGATA 1295
```

Search completed: January 8, 2006, 19:06:34
Job time : 993 secs

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 8, 2006, 17:06:59 ; Search time 1732 Seconds
(without alignments)
722.030 Million cell updates/sec

Title: US-10-028-415-20
Perfect score: 22
Sequence: 1 ctgcacaggagggttggaatac 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	22	100.0	250	6	AX322171	AX322171 Sequence
	2	22	100.0	400	6	BD059641	BD059641 Secreted
	3	22	100.0	418	10	BV196146	BV196146 sqnm18735
c	4	22	100.0	537	10	BV335238	BV335238 S230P6449
c	5	22	100.0	545	6	CQ922201	CQ922201 Sequence
c	6	22	100.0	571	10	BV312092	BV312092 S236P6311
c	7	22	100.0	775	6	CQ720809	CQ720809 Sequence
	8	22	100.0	822	6	AX284014	AX284014 Sequence
c	9	22	100.0	917	8	BC000064	BC000064 Homo sapi
c	10	22	100.0	922	8	BC072014	BC072014 Homo sapi
c	11	22	100.0	923	8	AB179457	AB179457 Macaca fa
c	12	22	100.0	954	6	AX417444	AX417444 Sequence
c	13	22	100.0	975	11	AY888736	AY888736 Synthetic
c	14	22	100.0	975	11	AY891136	AY891136 Synthetic
c	15	22	100.0	975	11	AY891391	AY891391 Synthetic
c	16	22	100.0	1452	8	HUMDBPB	M24070 Human DNA-b
c	17	22	100.0	1454	8	BC015208	BC015208 Homo sapi
c	18	22	100.0	1458	8	HUMAAE	L28809 Homo sapien
c	19	22	100.0	1468	8	HUMNSEP	M83234 Human nucle
c	20	22	100.0	1474	6	CS032279	CS032279 Sequence
c	21	22	100.0	1474	6	CS041231	CS041231 Sequence
c	22	22	100.0	1474	8	HUMRNABP	M85234 Human nucle
c	23	22	100.0	1475	8	BC098435	BC098435 Homo sapi
c	24	22	100.0	1476	6	CQ724685	CQ724685 Sequence
c	25	22	100.0	1481	6	AR117695	AR117695 Sequence
c	26	22	100.0	1481	6	AX409453	AX409453 Sequence
c	27	22	100.0	1481	6	AX498410	AX498410 Sequence
c	28	22	100.0	1481	6	AX840171	AX840171 Sequence
c	29	22	100.0	1481	8	HUMYB1A	J03827 Y box bindi
c	30	22	100.0	1503	4	OCU16821	U16821 Oryctolagus
c	31	22	100.0	1513	8	BC090038	BC090038 Homo sapi
c	32	22	100.0	1521	8	BC038384	BC038384 Homo sapi
c	33	22	100.0	1529	8	BC065571	BC065571 Homo sapi
c	34	22	100.0	1543	8	BC002411	BC002411 Homo sapi
c	35	22	100.0	1548	6	BD203713	BD203713 Human nuc
c	36	22	100.0	1548	6	AX014868	AX014868 Sequence
c	37	22	100.0	1554	8	BC010430	BC010430 Homo sapi
c	38	22	100.0	1574	8	BC070084	BC070084 Homo sapi
c	39	22	100.0	1599	8	BC071708	BC071708 Homo sapi
c	40	22	100.0	2168	8	HUMPSDBPB	L37516 Homo sapien
c	41	22	100.0	3073	6	AR083654	AR083654 Sequence
c	42	22	100.0	3073	6	I22491	I22491 Sequence 31
	43	22	100.0	99886	8	AL135841	AL135841 Human DNA
c	44	22	100.0	144902	14	AC016114	AC016114 Homo sapi
c	45	22	100.0	161835	8	AC098484	AC098484 Homo sapi

ALIGNMENTS

RESULT 1
AX322171/c
LOCUS AX322171 250 bp DNA linear PAT 07-JAN-2002
DEFINITION Sequence 44 from Patent EP1162276.
ACCESSION AX322171
VERSION AX322171.1 GI:18093236
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Kramer,M.D., Winter,H. and Reinartz,J.
TITLE Mrna molecules to be used as indicators of the functional and activational state of t-lymphocytes
JOURNAL Patent: EP 1162276-A 44 12-DEC-2001;
Lynx Therapeutics GmbH (DE)
FEATURES Location/Qualifiers
source 1. .250
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 100.0%; Score 22; DB 6; Length 250;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCACAGGAGGGTTGGAATAC 22
||||||||||||||||||||||
Db 125 CTGCACAGGAGGGTTGGAATAC 104

RESULT 2
BD059641
LOCUS BD059641 400 bp DNA linear PAT 27-AUG-2002
DEFINITION Secreted expressed sequence tags (sESTs).
ACCESSION BD059641
VERSION BD059641.1 GI:22605247
KEYWORDS JP 2001518793-A/1.

```
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 400)
  AUTHORS   Jacobs,K., Mccoy,J.M., Lavallie,E.R., Racie,L.A., Merberg,D.,
            Treacy,M., Spaulding,V. and Agostino,M.J.
  TITLE     Secreted expressed sequence tags (sESTs)
  JOURNAL   Patent: JP 2001518793-A 1 16-OCT-2001;
            GENETICS INSTITUTE INC
COMMENT     PN   JP 2001518793-A/1
            PD   16-OCT-2001
            PF   10-APR-1998 JP 1998543070
            PR   10-APR-1997 US    08/837312
            PI   KENNETH JACOBS,JOHN M MCCOY,EDWARD R LAVALLIE,LISA A RACIE, PI
              DAVID MERBERG,
            PI   MAURICE TREACY,VIKKI SPAULDING,MICHAEL J AGOSTINO PC
            C12N15/12,C12N5/10,C07K14/47,C12Q1/68,A61K38/17 CC   Strandedness:
            Double;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers.
FEATURES                Location/Qualifiers
     source             1. .400
                        /organism="Zea mays"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:4577"
ORIGIN

  Query Match             100.0%;  Score 22;  DB 6;  Length 400;
  Best Local Similarity   100.0%;  Pred. No. 0.77;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        200 CTGCACAGGAGGGTTGGAATAC 221


RESULT 3
BV196146
LOCUS       BV196146                418 bp    DNA     linear   STS 10-JUN-2004
DEFINITION  sqnm187359 Human DNA (Sequenom) Homo sapiens STS genomic, sequence
            tagged site.
ACCESSION   BV196146
VERSION     BV196146.1  GI:48040237
KEYWORDS    STS.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 418)
  AUTHORS   Nelson,R.M., Marnellos,G., Kammerer,S., Hoyal,C.R., Shi,M.M.,
            Cantor,C.R. and Braun,A.
  TITLE     Large-Scale Validation of Single Nucleotide Polymorphisms in Gene
            Regions
  JOURNAL   Genome Res. (2004) In press
COMMENT
            Contact: Andreas Braun
            Pharmaceuticals division
            Sequenom, Inc.
            3595 John Hopkins Court, San Diego, CA 92121, USA
            Tel: 18582029018
            Fax: 18582029020
            Email: abraun@sequenom.com
            Primer A: No primer sequence submitted
            Primer B: No primer sequence submitted
            STS size: 418.
FEATURES                Location/Qualifiers
     source             1. .418
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /clone_lib="Human DNA (Sequenom)"
     STS                <1. .>418
ORIGIN

  Query Match             100.0%;  Score 22;  DB 10;  Length 418;
  Best Local Similarity   100.0%;  Pred. No. 0.77;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        226 CTGCACAGGAGGGTTGGAATAC 247


RESULT 4
BV335238/c
LOCUS       BV335238                537 bp    DNA     linear   STS 27-JAN-2005
DEFINITION  S230P6449RD8.T0 Rottweiler Canis familiaris STS genomic, sequence
            tagged site.
ACCESSION   BV335238
VERSION     BV335238.1  GI:57535541
KEYWORDS    STS.
SOURCE      Canis familiaris (dog)
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
            Canis.
REFERENCE   1  (bases 1 to 537)
  AUTHORS   Lindblad-Toh,K.
  TITLE     The genome sequence of Canis familiaris
  JOURNAL   Unpublished (2004)
COMMENT
            Contact: Kerstin Lindblad-Toh
            Whitehead Institute for Biomedical Research, Center for Genome
            Research
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172521477
            Fax: 6172580903
            Email: kersli@genome.wi.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 537
            Protocol:
            WGS-discovery (WGS):
            Paired-end low-coverage whole genome shotgun reads were generated
            from 9 breeds
            (German Shepherd, Rottweiler, Bedlington Terrier, Beagle, Labrador
            Retriever, English
            Shepherd, Italian Greyhound, Alaskan Malamute and the Portuguese
            Water Dog -100,000 each)
            and five other canids (Chinese, Alaskan, Indian and Spanish Gray
            Wolf as well as the
            Californian Coyote).
            The WGS reads were placed uniquely on the CanFam1.0 boxer assembly
            and SNP detection was
            carried out by SSAHA-SNP. 863872 reads were annotated as STSs and
            485941 SNPs were
            annotated with alleles from the boxer and the breed or canid from
            which the particular
            read came. The validation rate for these SNPs was estimated at
            approximately 98%.
            WGA-discovery (WGA) of Boxer/Poodle SNPs:
            A second set of SNPs was generated using a similar methodology
            except that the contigs
            from the 1.5x poodle assembly (Kirkness 2003) were used instead of
            WGS reads. Since this
            sequence lacked base quality scores, arbitrary quality scores of
            phred 40 were assigned
            before the poddle seqeunce was placed uniquely on the CanFam1.0
            boxer assembly and SNP
            detection was carried out by SSAHA-SNP. 1637780 SNPs were annotated
            with alleles from the
            boxer and the poodle. The validation rate for these SNPs was
            estimated at approximately TBD%.
```

```
            Internal-WGA-discovery (I-WGA):
            A third set of SNPs were discovered by comparing reads in the WGA
            assembly. SNPs were
            defined as mismatch positions that had a base quality of >= 30 on
            both reads in a region
            that aligned without gaps, and with at most one additional mismatch
            in the ten flanking
            bases. For each allele, at least one additional read had to confirm
            it. 731476 SNPs were
            annotated with alleles between the two boxer alleles. The
            validation rate for these SNPs
            was estimated at approximately TBD%.
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Canis familiaris"
                     /mol_type="genomic DNA"
                     /strain="Rottweiler"
                     /db_xref="taxon:9615"
                     /map="+ 2 22-491 14601761-14601292"
                     /clone_lib="Rottweiler"
     STS             <1. .>537
ORIGIN

  Query Match            100.0%;  Score 22;  DB 10;  Length 537;
  Best Local Similarity  100.0%;  Pred. No. 0.77;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        514 CTGCACAGGAGGGTTGGAATAC 493


RESULT 5
CQ922201/c
LOCUS       CQ922201                545 bp    DNA     linear   PAT 23-NOV-2004
DEFINITION  Sequence 3401 from Patent WO2004097052.
ACCESSION   CQ922201
VERSION     CQ922201.1  GI:56212142
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Burczynski,M.E., Twine,N.C., Slonim,D.K., Trepicchio,W.L.,
            Strahs,A., Immerman,F. and Dorner,A.J.
  TITLE     Methods for prognosis and treatment of solid tumors
  JOURNAL   Patent: WO 2004097052-A 3401 11-NOV-2004;
            Wyeth (US); Burczynski, Michael E. (US)
FEATURES             Location/Qualifiers
     source          1. .545
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     misc_feature    41
                     /note="n includes a, c, g, or t, or contains no
                     nucleotide"
     misc_feature    79
                     /note="n includes a, c, g, or t, or contains no
                     nucleotide"
     misc_feature    114
                     /note="n includes a, c, g, or t, or contains no
                     nucleotide"
     misc_feature    122
                     /note="n includes a, c, g, or t, or contains no
                     nucleotide"
     misc_feature    131
                     /note="n includes a, c, g, or t, or contains no
                     nucleotide"
     misc_feature    134
                     /note="n includes a, c, g, or t, or contains no
                     nucleotide"
     misc_feature    138
                     /note="n includes a, c, g, or t, or contains no
                     nucleotide"
     misc_feature    231. .234
                     /note="n includes a, c, g, or t, or contains no
                     nucleotide"
     misc_feature    236. .237
                     /note="n includes a, c, g, or t, or contains no
                     nucleotide"
     misc_feature    239. .244
                     /note="n includes a, c, g, or t, or contains no
                     nucleotide"
     misc_feature    246. .254
                     /note="n includes a, c, g, or t, or contains no
                     nucleotide"
     misc_feature    256. .260
                     /note="n includes a, c, g, or t, or contains no
                     nucleotide"
     misc_feature    263
                     /note="n includes a, c, g, or t, or contains no
                     nucleotide"
ORIGIN

  Query Match            100.0%;  Score 22;  DB 6;  Length 545;
  Best Local Similarity  100.0%;  Pred. No. 0.77;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db         36 CTGCACAGGAGGGTTGGAATAC 15


RESULT 6
BV312092/c
LOCUS       BV312092                571 bp    DNA     linear   STS 26-JAN-2005
DEFINITION  S236P6311RA2.T0 AlaskanMalamute Canis familiaris STS genomic,
            sequence tagged site.
ACCESSION   BV312092
VERSION     BV312092.1  GI:57510584
KEYWORDS    STS.
SOURCE      Canis familiaris (dog)
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
            Canis.
REFERENCE   1  (bases 1 to 571)
  AUTHORS   Lindblad-Toh,K.
  TITLE     The genome sequence of Canis familiaris
  JOURNAL   Unpublished (2004)
COMMENT
            Contact: Kerstin Lindblad-Toh
            Whitehead Institute for Biomedical Research, Center for Genome
            Research
            320 Charles Street, Cambridge, MA 02141, USA
            Tel: 6172521477
            Fax: 6172580903
            Email: kersli@genome.wi.mit.edu
            Primer A: No sequence submitted
            Primer B: No sequence submitted
            STS size: 571
            Protocol:
            WGS-discovery (WGS):
            Paired-end low-coverage whole genome shotgun reads were generated
            from 9 breeds
            (German Shepherd, Rottweiler, Bedlington Terrier, Beagle, Labrador
            Retriever, English
            Shepherd, Italian Greyhound, Alaskan Malamute and the Portuguese
            Water Dog -100,000 each)
            and five other canids (Chinese, Alaskan, Indian and Spanish Gray
            Wolf as well as the
            Californian Coyote).
            The WGS reads were placed uniquely on the CanFam1.0 boxer assembly
            and SNP detection was
```

carried out by SSAHA-SNP. 863872 reads were annotated as STSs and 485941 SNPs were
annotated with alleles from the boxer and the breed or canid from which the particular
read came. The validation rate for these SNPs was estimated at approximately 98%.
WGA-discovery (WGA) of Boxer/Poodle SNPs:
A second set of SNPs was generated using a similar methodology except that the contigs
from the 1.5x poodle assembly (Kirkness 2003) were used instead of WGS reads. Since this
sequence lacked base quality scores, arbitrary quality scores of phred 40 were assigned
before the poddle seqeunce was placed uniquely on the CanFam1.0 boxer assembly and SNP
detection was carried out by SSAHA-SNP. 1637780 SNPs were annotated with alleles from the
boxer and the poodle. The validation rate for these SNPs was estimated at approximately TBD%.
Internal-WGA-discovery (I-WGA):
A third set of SNPs were discovered by comparing reads in the WGA assembly. SNPs were
defined as mismatch positions that had a base quality of >= 30 on both reads in a region
that aligned without gaps, and with at most one additional mismatch in the ten flanking
bases. For each allele, at least one additional read had to confirm it. 731476 SNPs were
annotated with alleles between the two boxer alleles. The validation rate for these SNPs
was estimated at approximately TBD%.

```
FEATURES             Location/Qualifiers
     source          1. .571
                     /organism="Canis familiaris"
                     /mol_type="genomic DNA"
                     /strain="AlaskanMalamute"
                     /db_xref="taxon:9615"
                     /map="+ 2 22-517 14601592-14601097"
                     /clone_lib="AlaskanMalamute"
     STS             <1. .>571
ORIGIN

  Query Match             100.0%;  Score 22;  DB 10;  Length 571;
  Best Local Similarity   100.0%;  Pred. No. 0.77;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        345 CTGCACAGGAGGGTTGGAATAC 324


RESULT 7
CQ720809/c
LOCUS       CQ720809                775 bp    DNA     linear   PAT 03-FEB-2004
DEFINITION  Sequence 6743 from Patent WO02068579.
ACCESSION   CQ720809
VERSION     CQ720809.1  GI:42281666
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
  TITLE     Kits, such as nucleic acid arrays, comprising a majority of
            humanexons or transcripts, for detecting expression and other uses
            thereof
  JOURNAL   Patent: WO 02068579-A 6743 06-SEP-2002;
            PE Corporation (NY) (US)
FEATURES             Location/Qualifiers
     source          1. .775
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             100.0%;  Score 22;  DB 6;  Length 775;
  Best Local Similarity   100.0%;  Pred. No. 0.77;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        442 CTGCACAGGAGGGTTGGAATAC 421


RESULT 8
AX284014
LOCUS       AX284014                822 bp    RNA     linear   PAT 20-NOV-2001
DEFINITION  Sequence 287 from Patent WO0179287.
ACCESSION   AX284014
VERSION     AX284014.1  GI:17044725
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Palin,M.F., Pomar,C. and gari Py,C.
  TITLE     Steatosis-modulating factors and uses thereof
  JOURNAL   Patent: WO 0179287-A 287 25-OCT-2001;
            Sa Majest la Reine du Chef du Canada Agriculture (CA);
            Agroalimentaire Canada (CA)
FEATURES             Location/Qualifiers
     source          1. .822
                     /organism="synthetic construct"
                     /mol_type="unassigned RNA"
                     /db_xref="taxon:32630"
                     /note="Expressed Sequence Tag Porcine Muscular steatosis"
ORIGIN

  Query Match             100.0%;  Score 22;  DB 6;  Length 822;
  Best Local Similarity   100.0%;  Pred. No. 0.77;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db         71 CTGCACAGGAGGGTTGGAATAC 92


RESULT 9
BC000064/c
LOCUS       BC000064                917 bp    mRNA    linear   PRI 30-SEP-2003
DEFINITION  Homo sapiens nuclease sensitive element binding protein 1, mRNA
            (cDNA clone IMAGE:3510030), partial cds.
ACCESSION   BC000064
VERSION     BC000064.2  GI:33875176
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 917)
  AUTHORS   Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S.,
```

Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCE 2 (bases 1 to 917)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (03-NOV-2000) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT On Aug 19, 2003 this sequence version replaced gi:13937530.
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbiology.org
contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 7 Row: p Column: 20.

FEATURES Location/Qualifiers
source 1. .917
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3510030"
/tissue_type="Brain, neuroblastoma"
/clone_lib="NIH_MGC_19"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
gene 1. .>917
/gene="NSEP1"
/note="synonyms: YB1, DBPB, YB-1, NSEP-1, MDR-NF1, BP-8, YBX1, CSDB"
/db_xref="GeneID:4904"
/db_xref="MIM:154030"
CDS 119. .>917
/gene="NSEP1"
/codon_start=1
/product="NSEP1 protein"
/protein_id="AAH00064.1"
/db_xref="GI:33875177"
/db_xref="GeneID:4904"
/db_xref="MIM:154030"
/translation="MSSEAETQQPPAAPPAAPALSAADTKPGTTGSGAGSGGPGGLTS
AAPAGGDKKVIATKVLGTVKWFNVRNGYGFINRNDTKEDVFVHQTAIKKNNPRKYLRS
VGDGETVEFDVVEGEKGAEAANVTGPGGVPVQGSKYAADRNHYRRYPRRRGPPRNYQQ
NYQNSESGEKNEGSESAPEGQAQQRRPYRRRRFPPYYMRRPYGRRPQYSNPPVQGEVM
EGADNQGAGEQGRPVRQNMYRGYRPRFRRGPPRQRQPREDGNEKKKKK"
misc_feature 290. .502
/gene="NSEP1"
/note="cold; Region: Cold Shock RNA binding domain of the OB fold"
/db_xref="CDD:LOAD_cold"

ORIGIN

Query Match 100.0%; Score 22; DB 8; Length 917;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCACAGGAGGGTTGGAATAC 22
|||||||||||||||||||||
Db 760 CTGCACAGGAGGGTTGGAATAC 739

RESULT 10
BC072014/c
LOCUS BC072014 922 bp mRNA linear PRI 03-JAN-2005
DEFINITION Homo sapiens nuclease sensitive element binding protein 1, mRNA (cDNA clone IMAGE:5556363), partial cds.
ACCESSION BC072014
VERSION BC072014.1 GI:47939142
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 922)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Ketteman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED 12477932
REFERENCE 2 (bases 1 to 922)
AUTHORS Director MGC Project.
TITLE Direct Submission
JOURNAL Submitted (01-JUN-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 58 Row: h Column: 23
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 34098945.

FEATURES Location/Qualifiers
source 1. .922
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5556363"
/tissue_type="Kidney, Wilm's tumor"
/clone_lib="NIH_MGC_108"

```
                     /lab_host="DH10B-R"
                     /note="Vector: pOTB7"
     gene            1. .>922
                     /gene="NSEP1"
                     /note="synonyms: DBPB, YB-1, NSEP-1, MDR-NF1, BP-8, YBX1,
                     CSDB, YB1"
                     /db_xref="GeneID:4904"
                     /db_xref="MIM:154030"
     CDS             124. .>922
                     /gene="NSEP1"
                     /codon_start=1
                     /product="NSEP1 protein"
                     /protein_id="AAH72014.1"
                     /db_xref="GI:47939143"
                     /db_xref="GeneID:4904"
                     /db_xref="MIM:154030"
                     /translation="MSSEAETQQPPAAPPAAPALSAADTKPGTTGSGAGSGGPGGLTS
                     AAPAGGDKKVIATKVLGTVKWFNVRNGYGFINRNDTKEDVFVHQTAIKKNNPRKYLRS
                     VGDGETVEFDVVEGEKGAEAANVTGPGGVPVQGSKYAADRNHYRRYPRRRGPPRNYQQ
                     NYQNSESGEKNEGSESAPEGQAQQRRPYRRRRFPPYYMRRPYGRRPQYSNPPVQGEVM
                     EGADNQGAGEQGRPVRQNMYRGYRPRFRRGPPRQRQPREDGNEKKKKK"
ORIGIN

  Query Match           100.0%;  Score 22;  DB 8;  Length 922;
  Best Local Similarity  100.0%;  Pred. No. 0.77;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        765 CTGCACAGGAGGGTTGGAATAC 744


RESULT 11
AB179457/c
LOCUS       AB179457                923 bp    mRNA    linear   PRI 18-JUN-2005
DEFINITION  Macaca fascicularis testis cDNA clone: QtsA-20374, similar to human
            nuclease sensitive element binding protein 1 (NSEP1), mRNA, RefSeq:
            NM_004559.2.
ACCESSION   AB179457
VERSION     AB179457.1  GI:67972331
KEYWORDS    oligo capping; fis (full insert sequence).
SOURCE      Macaca fascicularis (crab-eating macaque)
  ORGANISM  Macaca fascicularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Cercopithecidae; Cercopithecinae; Macaca.
REFERENCE   1
  AUTHORS   International consortium for macaque cDNA sequencing and analysis.
  TITLE     DNA sequences of macaque genes expressed in brain or testis and its
            evolutionary implications
  JOURNAL   Unpublished
REFERENCE   2
  AUTHORS   Osada,N., Hirata,M., Tanuma,R., Kusuda,J., Hida,M., Suzuki,Y.,
            Sugano,S., Gojobori,T., Shen,J.C.-K., Wu,C.I. and Hashimoto,K.
  TITLE     Substitution rate and structural divergence of 5'UTR evolution:
            Comparative analysis between human and cynomolgus monkey cDNAs
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 923)
  AUTHORS   Hashimoto,K., Kusuda,J. and Sugano,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-MAY-2004) Katsuyuki Hashimoto, National Institute of
            Infectious Diseases, Division of Genetic Resources; 23-1, Toyama
            1-chome, Shinjuku-ku, Tokyo, 162-8640, Japan
            (E-mail:khashi@nih.go.jp, URL:http://www.nih.go.jp/yoken/genebank/,
            Tel:81-3-5285-1111(ex.2120), Fax:81-3-5285-1181)
COMMENT     The International consortium for macaque cDNA sequencing and
            analysis consists of:Department of Virology and Human Genome
            Center, Institute of Medical Science, The University of Tokyo,
            Tokyo, Japan; Division of Genetic Resources, National Institute of
            Infectious Diseases of Japan, Tokyo, Japan; National Health
            Research Institute, Taipei, Taiwan; Institute of Molecular Biology,
            Academia Sinica, Taipei, Taiwan; Department of Ecology & Evolution,
            University of Chicago, Chicago, IL, USA, Center for Information
            Biology,
            National Institute of Genetics of Japan, Mishima, Japan. Clone
            distribution: clone distribution information can be found at:
            http://www.nih.go.jp/yoken/genebank/
            Lab host:    TOP10
            Vector:      pME18S-FL3 (Acc.No. AB009864)
            R.  Site1:   DraIII (CACTGTGTG)
            R.  Site2:   DraIII (CACCATGTG)
            Description: 1st strand cDNA was primed with an oligo(dT) primer
            [ATGTGGCCTTTTTTTTTTTTTTTT]; double-stranded cDNA was synthesized
            using specific 5' and 3' primers and amplified by PCR.  The PCR
            product was digested with SfiI and size selection was performed to
            exclude fragments <1.5kb.The SfiI-digested PCR product was cloned
            into distinct DraIII sites of pME18S-FL3. XhoI sites just outside
            the DraIII sites can be used to isolate the cDNA insert.  Libraries
            were constructed by oligo-capping method. Libraries were made from:
                         QccE: cerebellum cortex
                         QnpA: parietal lobe
                         QtrA: temporal lobe right
                         QflA: frontal lobe left
                         QmoA: medulla oblongata
                         QbsA: brain stem
                         QorA: occipital lobe right
                         QtsA: testis
            Custom primers were used for 5' and 3'-end sequencing. The
            full-insert
            sequencing was done by primer-walking method using ABI DNA
            sequencer.
FEATURES             Location/Qualifiers
     source          1. .923
                     /organism="Macaca fascicularis"
                     /mol_type="mRNA"
                     /db_xref="taxon:9541"
                     /clone="QtsA-20374"
                     /sex="male"
                     /clone_lib="macaque cDNA library QtsA"
                     /dev_stage="adult"
     CDS             91. .474
                     /note="unnamed protein product; Homo sapiens nuclease
                     sensitive element binding protein 1 (NSEP1), mRNA, RefSeq:
                     NM_004559.2"
                     /codon_start=1
                     /protein_id="BAE02508.1"
                     /db_xref="GI:67972332"
                     /translation="MRRPYGRRPQYSNPPVQGEVMEGADNQGAGEQGRPVRQNMYRGY
                     RPRFRRGPPRQRQPREDGNEEDKENQGDETQGQQPPQRRYRRNFNYRRRRPENPKPQD
                     GKETKAADPPAENSSAPEAEQGGAE"
ORIGIN

  Query Match           100.0%;  Score 22;  DB 8;  Length 923;
  Best Local Similarity  100.0%;  Pred. No. 0.77;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        141 CTGCACAGGAGGGTTGGAATAC 120


RESULT 12
AX417444/c
LOCUS       AX417444                954 bp    DNA     linear   PAT 18-JUN-2002
DEFINITION  Sequence 26 from Patent EP1197495.
ACCESSION   AX417444
VERSION     AX417444.1  GI:21522729
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1
  AUTHORS   Higashi,K. and Komatsu,K.
```

```
  TITLE     Dna-binding protein yb-1-containing collagen accumulation
            inhibitors
  JOURNAL   Patent: EP 1197495-A 26 17-APR-2002;
            Sumitomo Chemical Company, Limited (JP)
FEATURES             Location/Qualifiers
     source          1. .954
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             1. .954
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD35407.1"
                     /db_xref="GI:21522730"
                     /translation="MSSEAETQQPPAAPPAAPALSAADTKPGTTGSGAGSGGPGGLTS
                     AAPAGGDKKVIATKVLGTVKWFNVRNGYGFINRNDTKEDVFVHQTAIKKNNPRKYLRS
                     VGDGETVEFDVVEGEKGEEAANVTGPGGVPVQGSKYAADRNHYRRYPRRRGPPRNYQQ
                     NYQNSESGEKNEGSESAPEGQAQQRRPYRRRRFPPYYMRRPYGRRPQYSNPPVQGEVM
                     EGADNQGAGEQGRPVRQNMYRGYRPRFRRGPPRQRQPREDGNEEDKENQGDETQGQQP
                     PQRRYRRNFNYRRRRPENPKPQDGKETKAADPPAENSRSRG"
ORIGIN

  Query Match             100.0%;  Score 22;  DB 6;  Length 954;
  Best Local Similarity   100.0%;  Pred. No. 0.77;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        642 CTGCACAGGAGGGTTGGAATAC 621


RESULT 13
AY888736/c
LOCUS       AY888736                975 bp    mRNA    linear   SYN 29-MAR-2005
DEFINITION  Synthetic construct Homo sapiens clone FLH030953.01X nuclease
            sensitive element binding protein 1 (NSEP1) mRNA, complete cds.
ACCESSION   AY888736
VERSION     AY888736.1  GI:61359168
KEYWORDS    Human ORF Project.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1  (bases 1 to 975)
  AUTHORS   Hines,L., Rolfs,A., Jepson,D., Moreira,D., Raphael,J., Kelley,F.,
            Shen,B., Halleck,A., Koundinya,M., Hu,Y., Zuo,D., Taycher,E.,
            Williamson,J. and LaBaer,J.
  TITLE     Cloning of human full-length CDS in Creator (TM) recombinational
            vector system
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 975)
  AUTHORS   Hines,L., Rolfs,A., Jepson,D., Moreira,D., Raphael,J., Kelley,F.,
            Shen,B., Halleck,A., Koundinya,M., Hu,Y., Zuo,D., Taycher,E.,
            Williamson,J. and LaBaer,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-JAN-2005) Biological Chemistry and Molecular
            Pharmacology, Harvard Institute of Proteomics, 320 Charles St.,
            Cambridge, MA 02141, USA
COMMENT     This CDS clone is a part of a collection of human full-length
            expression clones generated by Harvard Institute of Proteomics.
            This ORF clone has been cloned with normalized stop-codon. The CDS
            has been directionally cloned using BD In-Fusion(TM) cloning system
            between the SalI and HindIII sites of the pDNR-Dual vector.
            Additional sequences in the clone:  'ACC' after SalI site and
            before 'ATG' to provide Kozak consensus sequence.  Each clone is
            clonally isolated and full-length sequence-verified.
FEATURES             Location/Qualifiers
     source          1. .975
                     /organism="synthetic construct"
                     /mol_type="mRNA"
                     /db_xref="taxon:32630"
                     /clone="FLH030953.01X"
                     /lab_host="Escherichia coli DH5alpha T1 resistant"
                     /note="derived from MGC template"
     gene            1. .975
                     /gene="NSEP1"
     CDS             1. .975
                     /gene="NSEP1"
                     /codon_start=1
                     /transl_table=11
                     /product="nuclease sensitive element binding protein 1"
                     /protein_id="AAX41678.1"
                     /db_xref="GI:61359169"
                     /translation="MSSEAETQQPPAAPPAAPALSAADTKPGTTGSGAGSGGPGGLTS
                     AAPAGGDKKVIATKVLGTVKWFNVRNGYGFINRNDTKEDVFVHQTAIKKNNPRKYLRS
                     VGDGETVEFDVVEGEKGAEAANVTGPGGVPVQGSKYAADRNHYRRYPRRRGPPRNYQQ
                     NYQNSESGEKNEGSESAPEGQAQQRRPYRRRRFPPYYMRRPYGRRPQYSNPPVQGEVM
                     EGADNQGAGEQGRPVRQNMYRGYRPRFRRGPPRQRQPREDGNEEDKENQGDETQGQQP
                     PQRRYRRNFNYRRRRPENPKPQDGKETKAADPPAENSSAPEAEQGGAE"
ORIGIN

  Query Match             100.0%;  Score 22;  DB 11;  Length 975;
  Best Local Similarity   100.0%;  Pred. No. 0.77;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        642 CTGCACAGGAGGGTTGGAATAC 621


RESULT 14
AY891136/c
LOCUS       AY891136                975 bp    mRNA    linear   SYN 21-MAR-2005
DEFINITION  Synthetic construct Homo sapiens clone FLH007972.01L nuclease
            sensitive element binding protein 1 (NSEP1) mRNA, partial cds.
ACCESSION   AY891136
VERSION     AY891136.1  GI:61367906
KEYWORDS    Human ORF Project.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1  (bases 1 to 975)
  AUTHORS   Hines,L., Rolfs,A., Jepson,D., Moreira,D., Raphael,J., Kelley,F.,
            Shen,B., Halleck,A., Koundinya,M., Hu,Y., Zuo,D., Taycher,E.,
            Williamson,J. and LaBaer,J.
  TITLE     Cloning of human full-length CDS in Creator (TM) recombinational
            vector system
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 975)
  AUTHORS   Hines,L., Rolfs,A., Jepson,D., Moreira,D., Raphael,J., Kelley,F.,
            Shen,B., Halleck,A., Koundinya,M., Hu,Y., Zuo,D., Taycher,E.,
            Williamson,J. and LaBaer,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JAN-2005) Biological Chemistry and Molecular
            Pharmacology, Harvard Institute of Proteomics, 320 Charles St.,
            Cambridge, MA 02141, USA
COMMENT     This CDS clone is a part of a collection of human full-length
            expression clones generated by Harvard Institute of Proteomics.
            This ORF clone has been cloned without stop-codon (to allow fusion
            with C-terminal tag). The CDS has been directionally cloned using
            BD In-Fusion(TM) cloning system between the SalI and HindIII sites
            of the pDNR-Dual vector. Additional sequences in the clone:  'ACC'
            after SalI site and before 'ATG' to provide Kozak consensus
            sequence; 'GG' after last codon and before HindIII site to maintain
            reading frame.  Each clone is clonally isolated and full-length
            sequence-verified.
FEATURES             Location/Qualifiers
     source          1. .975
                     /organism="synthetic construct"
                     /mol_type="mRNA"
                     /db_xref="taxon:32630"
                     /clone="FLH007972.01L"
                     /lab_host="Escherichia coli DH5alpha T1 resistant"
                     /note="derived from Homo sapiens first strand cDNA library
                     from placenta and brain"
     gene            1. .>975
                     /gene="NSEP1"
```

```
     CDS             1. .>975
                     /gene="NSEP1"
                     /codon_start=1
                     /transl_table=11
                     /product="nuclease sensitive element binding protein 1"
                     /protein_id="AAX43065.1"
                     /db_xref="GI:61367907"
                     /translation="MSSEAETQQPPAAPPAAPALSAADTKPGTTGSGAGSGGPGGLTS
                     AAPAGGDKKVIATKVLGTVKWFNVRNGYGFINRNDTKEDVFVHQTAIKKNNPRKYLRS
                     VGDGETVEFDVVEGEKGAEAANVTGPGGVPVQGSKYAADRNHYRRYPRRRGPPRNYQQ
                     NYQNSESGEKNEGSESAPEGQAQQRRPYRRRRFPPYYMRRPYGRRPQYSNPPVQGEVM
                     EGADNQGAGEQGRPVRQNMYRGYRPRFRRGPPRQRQPREDGNEEDKENQGDETQGQQP
                     PQRRYRRNFNYRRRRPENPKPQDGKETKAADPPAENSSAPEAEQGGAEL"
ORIGIN

  Query Match           100.0%;  Score 22;  DB 11;  Length 975;
  Best Local Similarity  100.0%;  Pred. No. 0.77;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        642 CTGCACAGGAGGGTTGGAATAC 621


RESULT 15
AY891391/c
LOCUS       AY891391                975 bp    mRNA    linear   SYN 29-MAR-2005
DEFINITION  Synthetic construct Homo sapiens clone FLH030949.01L nuclease
            sensitive element binding protein 1 (NSEP1) mRNA, partial cds.
ACCESSION   AY891391
VERSION     AY891391.1  GI:61369329
KEYWORDS    Human ORF Project.
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1  (bases 1 to 975)
  AUTHORS   Hines,L., Rolfs,A., Jepson,D., Moreira,D., Raphael,J., Kelley,F.,
            Shen,B., Halleck,A., Koundinya,M., Hu,Y., Zuo,D., Taycher,E.,
            Williamson,J. and LaBaer,J.
  TITLE     Cloning of human full-length CDS in Creator (TM) recombinational
            vector system
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 975)
  AUTHORS   Hines,L., Rolfs,A., Jepson,D., Moreira,D., Raphael,J., Kelley,F.,
            Shen,B., Halleck,A., Koundinya,M., Hu,Y., Zuo,D., Taycher,E.,
            Williamson,J. and LaBaer,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JAN-2005) Biological Chemistry and Molecular
            Pharmacology, Harvard Institute of Proteomics, 320 Charles St.,
            Cambridge, MA 02141, USA
COMMENT     This CDS clone is a part of a collection of human full-length
            expression clones generated by Harvard Institute of Proteomics.
            This ORF clone has been cloned without stop-codon (to allow fusion
            with C-terminal tag). The CDS has been directionally cloned using
            BD In-Fusion(TM) cloning system between the SalI and HindIII sites
            of the pDNR-Dual vector. Additional sequences in the clone:  'ACC'
            after SalI site and before 'ATG' to provide Kozak consensus
            sequence; 'GG' after last codon and before HindIII site to maintain
            reading frame.  Each clone is clonally isolated and full-length
            sequence-verified.
FEATURES             Location/Qualifiers
     source          1. .975
                     /organism="synthetic construct"
                     /mol_type="mRNA"
                     /db_xref="taxon:32630"
                     /clone="FLH030949.01L"
                     /lab_host="Escherichia coli DH5alpha T1 resistant"
                     /note="derived from MGC template"
     gene            1. .>975
                     /gene="NSEP1"
     CDS             1. .>975
                     /gene="NSEP1"
                     /codon_start=1
                     /transl_table=11
                     /product="nuclease sensitive element binding protein 1"
                     /protein_id="AAX43318.1"
                     /db_xref="GI:61369330"
                     /translation="MSSEAETQQPPAAPPAAPALSAADTKPGTTGSGAGSGGPGGLTS
                     AAPAGGDKKVIATKVLGTVKWFNVRNGYGFINRNDTKEDVFVHQTAIKKNNPRKYLRS
                     VGDGETVEFDVVEGEKGAEAANVTGPGGVPVQGSKYAADRNHYRRYPRRRGPPRNYQQ
                     NYQNSESGEKNEGSESAPEGQAQQRRPYRRRRFPPYYMRRPYGRRPQYSNPPVQGEVM
                     EGADNQGAGEQGRPVRQNMYRGYRPRFRRGPPRQRQPREDGNEEDKENQGDETQGQQP
                     PQRRYRRNFNYRRRRPENPKPQDGKETKAADPPAENSSAPEAEQGGAEL"
ORIGIN

  Query Match           100.0%;  Score 22;  DB 11;  Length 975;
  Best Local Similarity  100.0%;  Pred. No. 0.77;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        642 CTGCACAGGAGGGTTGGAATAC 621


Search completed: January  8, 2006, 17:56:30
Job time : 1736 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 8, 2006, 17:08:45 ; Search time 2518 Seconds
(without alignments)
408.783 Million cell updates/sec

Title: US-10-028-415-20
Perfect score: 22
Sequence: 1 ctgcacaggagggttggaatac 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	22	100.0	95	3	BM194709	BM194709 EST356 Sw
c 2	22	100.0	132	1	AI905848	AI905848 QV-BT100-
c 3	22	100.0	137	8	N85328	N85328 J3215F Huma
4	22	100.0	152	1	AI905850	AI905850 QV-BT100-
c 5	22	100.0	178	2	BF757327	BF757327 MR0-CT054
6	22	100.0	181	8	T63887	T63887 yc16h10.s1
c 7	22	100.0	199	3	BM851169	BM851169 K-EST0131
c 8	22	100.0	205	1	AW361577	AW361577 PM0-CT026
9	22	100.0	211	1	AI907390	AI907390 QV-BT141-
10	22	100.0	218	2	BF827899	BF827899 RC0-HN002
c 11	22	100.0	229	3	BQ048903	BQ048903 AGENCOURT
c 12	22	100.0	232	3	BM849989	BM849989 K-EST0130
c 13	22	100.0	239	1	AI905842	AI905842 IL-BT100-
14	22	100.0	264	1	AW059746	AW059746 LE3a11.yg
c 15	22	100.0	270	2	BG990119	BG990119 MR3-HT110
16	22	100.0	276	2	BF858412	BF858412 RC1-FT019
17	22	100.0	281	2	BF991764	BF991764 MR0-GN013
c 18	22	100.0	287	2	BF331397	BF331397 IL5-BT057
c 19	22	100.0	287	3	BM013541	BM013541 603639012
20	22	100.0	289	2	BF991767	BF991767 MR0-GN013
21	22	100.0	294	2	BF991360	BF991360 MR0-GN013
c 22	22	100.0	295	8	W16742	W16742 zb17c11.r1
c 23	22	100.0	300	2	BF762156	BF762156 PM4-CS003
c 24	22	100.0	304	2	BF378675	BF378675 MR2-UM006
c 25	22	100.0	304	8	T29674	T29674 EST89748 Hu
c 26	22	100.0	305	8	N46380	N46380 yy40c03.r1
c 27	22	100.0	325	1	AW889366	AW889366 QV4-NT002
c 28	22	100.0	328	3	BM850970	BM850970 K-EST0131
29	22	100.0	330	2	BF916151	BF916151 RC1-UT008
c 30	22	100.0	331	1	AI907387	AI907387 QV-BT141-
c 31	22	100.0	331	2	BE939753	BE939753 CM4-UT001
32	22	100.0	341	1	AW938845	AW938845 PM0-DT006
c 33	22	100.0	341	3	BM797655	BM797655 K-EST0080
c 34	22	100.0	343	1	AA172062	AA172062 zo98a04.r
c 35	22	100.0	343	6	CB145305	CB145305 K-EST0200
c 36	22	100.0	345	8	R01362	R01362 ye81e01.r1
c 37	22	100.0	347	2	BE843807	BE843807 RC0-TN007
38	22	100.0	349	1	AW068164	AW068164 cn23h08.x
39	22	100.0	357	1	AW804381	AW804381 QV0-UM009
c 40	22	100.0	357	2	BE825739	BE825739 QV3-EN000
41	22	100.0	361	2	BF991791	BF991791 MR0-GN013
42	22	100.0	362	2	BF991367	BF991367 MR0-GN013
c 43	22	100.0	362	8	D56380	D56380 HUM426F01B
c 44	22	100.0	364	1	AI908098	AI908098 RC-BT163-
c 45	22	100.0	364	2	BF858472	BF858472 RC1-FT019

ALIGNMENTS

RESULT 1
BM194709/c
LOCUS BM194709 95 bp mRNA linear EST 13-DEC-2001
DEFINITION EST356 Swine Xiang Longissimus Dorsi Muscle cDNA Library Sus scrofa cDNA clone pM356 5', mRNA sequence.
ACCESSION BM194709
VERSION BM194709.1 GI:17653695
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 95)
AUTHORS Wang,X., Li,N., Wu,C. and Li,C.
TITLE The Study on Pig Muscle Developmental Biology by Analysis of Expressed Sequence Tags
JOURNAL Unpublished (2001)
COMMENT Contact: Wang Xiuli
Department of Animal Science and Technology, The national key laboratory of agro-department
China Agricultural University
Yuan Ming Yuan West Road 2, Beijing, China, 100094
Email: xiuliwang318@263.net
This sequence may serve as a candidate gene of meat quality
Seq primer: T3 Primer.
FEATURES Location/Qualifiers
source 1. .95
/organism="Sus scrofa"
/mol_type="mRNA"
/strain="Xiang Pig"
/db_xref="taxon:9823"
/clone="pM356"
/tissue_type="Muscle"
/clone_lib="Swine Xiang Longissimus Dorsi Muscle cDNA Library"
/note="Organ: Longissimus Dorsi Muscle; The EST sequences were gotten from sequencing xiang pig muscle cDNA Library "
ORIGIN

Query Match 100.0%; Score 22; DB 3; Length 95;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db         64 CTGCACAGGAGGGTTGGAATAC 43


RESULT 2
AI905848/c
LOCUS       AI905848                132 bp    mRNA    linear   EST 30-MAR-2000
DEFINITION  QV-BT100-150399-021 BT100 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AI905848
VERSION     AI905848.1  GI:6496235
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 132)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/seq/gethtml.pl?t1=QV&t2=QV-BT100-021.html
            &t3=150399&t4=1)
            Seq primer: puc 18 forward.
FEATURES             Location/Qualifiers
     source          1. .132
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /sex="female"
                     /dev_stage="Adult"
                     /clone_lib="BT100"
                     /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No.
                     196,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
ORIGIN

  Query Match             100.0%;  Score 22;  DB 1;  Length 132;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db         46 CTGCACAGGAGGGTTGGAATAC 25


RESULT 3
N85328/c
LOCUS       N85328                  137 bp    mRNA    linear   EST 01-APR-1996
DEFINITION  J3215F Human fetal heart, Lambda ZAP Express Homo sapiens cDNA
            clone J3215 5' similar to DNA-BINDING PROTEIN NUCL.-SENSIT, mRNA
            sequence.
ACCESSION   N85328
VERSION     N85328.1  GI:1260953
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 137)
  AUTHORS   Liew,C.C.
  TITLE     cDNAs from fetal heart (1996)
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Liew CC
            Brigham and Women's Hospital
            Harvard Medical School
            75 Francis St. Boston, MA 02115, USA
            Tel: 6177328915
            Fax: 6179750995
            Email: cliew@rics.bwh.harvard.edu
            Seq primer: GAAATTAACCCTCACTAAAGGG.
FEATURES             Location/Qualifiers
     source          1. .137
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="J3215"
                     /lab_host="E. coli XL1-Blue"
                     /clone_lib="Human fetal heart, Lambda ZAP Express"
                     /note="Vector: Lambda ZAP Express; Site_1: EcoRI; Site_2:
                     XhoI; mRNA was purified from human fetal hearts (8-10
                     weeks).  cDNA was synthesized using a XhoI-Oligo dT
                     adaptor-primer.  EcoRI adaptors were ligated, followed by
                     digestion with XhoI, for directional cloning into
                     predigested lambda ZAP Express."
ORIGIN

  Query Match             100.0%;  Score 22;  DB 8;  Length 137;
  Best Local Similarity   100.0%;  Pred. No. 10;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db         35 CTGCACAGGAGGGTTGGAATAC 14


RESULT 4
AI905850
LOCUS       AI905850                152 bp    mRNA    linear   EST 30-MAR-2000
DEFINITION  QV-BT100-150399-017_1 BT100 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AI905850
VERSION     AI905850.1  GI:6496237
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 152)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
```

```
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/seq/gethtml.pl?t1=QV&t2=QV-BT100-017_1.ht
            ml&t3=150399&t4=1)
            Seq primer: puc 18 forward.
FEATURES              Location/Qualifiers
     source           1. .152
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /sex="female"
                      /dev_stage="Adult"
                      /clone_lib="BT100"
                      /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                      SmaI; A mini-library was made by cloning products derived
                      from  ORESTES PCR (U.S. Letters Patent application No.
                      196,716 -  Ludwig Institute for Cancer Research) profiles
                      into  the pUC 18 vector. Reverse transcription of tissue
                      mRNA and cDNA amplification were performed under low
                      stringency conditions."
ORIGIN

  Query Match             100.0%;  Score 22;  DB 1;  Length 152;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db         87 CTGCACAGGAGGGTTGGAATAC 108


RESULT 5
BF757327/c
LOCUS       BF757327                178 bp    mRNA    linear   EST 12-JAN-2001
DEFINITION  MR0-CT0540-041100-101-f02 CT0540 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF757327
VERSION     BF757327.1  GI:12105227
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 178)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR0&t2=MR0-CT0540-
            041100-101-f02&t3=2000-11-04&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 12
            High quality sequence stop: 178.
FEATURES              Location/Qualifiers
     source           1. .178
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /dev_stage="Adult"
                      /clone_lib="CT0540"
                      /note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
                      SmaI; A mini-library was made by cloning products derived
                      from  ORESTES PCR (U.S. Letters Patent application No.
                      196,716 -  Ludwig Institute for Cancer Research) profiles
                      into  the pUC 18 vector. Reverse transcription of tissue
                      mRNA and cDNA amplification were performed under low
                      stringency conditions."
ORIGIN

  Query Match             100.0%;  Score 22;  DB 2;  Length 178;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db         94 CTGCACAGGAGGGTTGGAATAC 73


RESULT 6
T63887
LOCUS       T63887                  181 bp    mRNA    linear   EST 17-FEB-1995
DEFINITION  yc16h10.s1 Stratagene lung (#937210) Homo sapiens cDNA clone
            IMAGE:80899 3' similar to gb:J03827 Y BOX BINDING PROTEIN-1
            (HUMAN);, mRNA sequence.
ACCESSION   T63887
VERSION     T63887.1  GI:667752
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 181)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W.,
            Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
            Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
            Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J.,
            Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.
            and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
   PUBMED   8889549
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 563
            High qality sequence stops: 149 Source: IMAGE Consortium, LLNL This
            clone is available royalty-free through LLNL ; contact the IMAGE
            Consortium (info@image.llnl.gov) for further information.
            Insert Length: 563   Std Error: 0.00
            Seq primer: -21m13
            High quality sequence stop: 149.
FEATURES              Location/Qualifiers
     source           1. .181
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="GDB:484516"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:80899"
                      /sex="male"
                      /dev_stage="72 years"
                      /lab_host="SOLR cells (kanamycin resistant)"
                      /clone_lib="Stratagene lung (#937210)"
```

```
                     /note="Organ: lung; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally.  Primer:
                     Oligo dT. normal lung.  Average insert size: 1.0 kb;
                     Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG
                     3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
ORIGIN

  Query Match             100.0%;  Score 22;  DB 8;  Length 181;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        112 CTGCACAGGAGGGTTGGAATAC 133


RESULT 7
BM851169/c
LOCUS       BM851169                 199 bp    mRNA    linear   EST 06-MAR-2002
DEFINITION  K-EST0131931 S19N665307 Homo sapiens cDNA clone S19N665307-17-A01
            5', mRNA sequence.
ACCESSION   BM851169
VERSION     BM851169.1  GI:19207568
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 199)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 17  row: A  column: 01
            High quality sequence stop: 199.
FEATURES             Location/Qualifiers
     source          1. .199
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S19N665307-17-A01"
                     /sex="M"
                     /lab_host="Top10F'"
                     /clone_lib="S19N665307"
                     /note="Organ: Stomach; Vector: pCNS; Site_1: EcoRI;
                     Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline phosphatase (BAP) and then decapped
                     with tabacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including EcoR
                     I site by treatment of T4 RNA ligase and the first strand
                     cDNA was synthesized from oligo dT-selected mRNA by
                     priming with dT-tailed vector. The dT-tailed vector was
                     adjusted to have about 60nt. The cDNA vector was
                     circularized with E. coli DNA ligase after digestion of
                     EcoRI which site is also included in vector. An RNA strand
                     converted to a DNA strand by Okayama-Berg method. The
                     obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation method.
                     The cDNA libraries constructed by this method are
                     full-length enriched cDNA library."
ORIGIN

  Query Match             100.0%;  Score 22;  DB 3;  Length 199;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        172 CTGCACAGGAGGGTTGGAATAC 151


RESULT 8
AW361577/c
LOCUS       AW361577                 205 bp    mRNA    linear   EST 04-FEB-2000
DEFINITION  PM0-CT0263-151099-002-e10 CT0263 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AW361577
VERSION     AW361577.1  GI:6866331
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 205)
  AUTHORS   HCGP http://www.ludwig.org.br/ORESTES.
  TITLE     The FAPESP/LICR Human Cancer Genome Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM0&t2=PM0-CT0263-
            151099-002-e10&t3=1999-10-15&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 8
            High quality sequence stop: 84.
FEATURES             Location/Qualifiers
     source          1. .205
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="CT0263"
                     /note="Organ: colon; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No.
                     196,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
ORIGIN

  Query Match             100.0%;  Score 22;  DB 1;  Length 205;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        123 CTGCACAGGAGGGTTGGAATAC 102


RESULT 9
AI907390
LOCUS       AI907390                 211 bp    mRNA    linear   EST 30-MAR-2000
DEFINITION  QV-BT141-180399-018 BT141 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AI907390
VERSION     AI907390.1  GI:6497820
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 211)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/seq/gethtml.pl?t1=QV&t2=QV-BT141-018.html
            &t3=180399&t4=1)
            Seq primer: puc 18 forward.
FEATURES             Location/Qualifiers
     source          1. .211
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /sex="female"
                     /dev_stage="Adult"
                     /clone_lib="BT141"
                     /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No.
                     196,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
ORIGIN

  Query Match             100.0%;  Score 22;  DB 1;  Length 211;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db         61 CTGCACAGGAGGGTTGGAATAC 82


RESULT 10
BF827899
LOCUS       BF827899                 218 bp    mRNA    linear   EST 13-JAN-2001
DEFINITION  RC0-HN0024-061200-031-e10 HN0024 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF827899
VERSION     BF827899.1  GI:12172185
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 218)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC0&t2=RC0-HN0024-
            061200-031-e10&t3=2000-12-06&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 58
            High quality sequence stop: 218.
FEATURES             Location/Qualifiers
     source          1. .218
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="HN0024"
                     /note="Organ: head_normal; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

  Query Match             100.0%;  Score 22;  DB 2;  Length 218;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        195 CTGCACAGGAGGGTTGGAATAC 216


RESULT 11
BQ048903/c
LOCUS       BQ048903                 229 bp    mRNA    linear   EST 29-MAR-2002
DEFINITION  AGENCOURT_6832553 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5789593
            5', mRNA sequence.
ACCESSION   BQ048903
VERSION     BQ048903.1  GI:19808243
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 229)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12883  row: g  column: 02
            High quality sequence stop: 228.
FEATURES             Location/Qualifiers
     source          1. .229
                     /organism="Homo sapiens"
```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5789593"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_92"
/note="Organ: testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 22; DB 3; Length 229;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCACAGGAGGGTTGGAATAC 22
||||||||||||||||||||||
Db 71 CTGCACAGGAGGGTTGGAATAC 50

RESULT 12
BM849989/c
LOCUS BM849989 232 bp mRNA linear EST 06-MAR-2002
DEFINITION K-EST0130517 S19N665307 Homo sapiens cDNA clone S19N665307-15-C01 5', mRNA sequence.
ACCESSION BM849989
VERSION BM849989.1 GI:19206388
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 232)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 15 row: C column: 01
High quality sequence stop: 232.
FEATURES Location/Qualifiers
source 1. .232
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S19N665307-15-C01"
/sex="M"
/lab_host="Top10F'"
/clone_lib="S19N665307"
/note="Organ: Stomach; Vector: pCNS; Site_1: EcoRI; Site_2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tabacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including EcoR I site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN

Query Match 100.0%; Score 22; DB 3; Length 232;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCACAGGAGGGTTGGAATAC 22
||||||||||||||||||||||
Db 172 CTGCACAGGAGGGTTGGAATAC 151

RESULT 13
AI905842/c
LOCUS AI905842 239 bp mRNA linear EST 30-MAR-2000
DEFINITION IL-BT100-301298-008 BT100 Homo sapiens cDNA, mRNA sequence.
ACCESSION AI905842
VERSION AI905842.1 GI:6496229
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 239)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/seq/gethtml.pl?t1=IL&t2=IL-BT100-008.html&t3=301298&t4=1)
Seq primer: puc 18 forward.
FEATURES Location/Qualifiers
source 1. .239
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/sex="female"
/dev_stage="Adult"
/clone_lib="BT100"
/note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 100.0%; Score 22; DB 1; Length 239;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCACAGGAGGGTTGGAATAC 22
||||||||||||||||||||||
Db 63 CTGCACAGGAGGGTTGGAATAC 42

```
RESULT 14
AW059746
LOCUS       AW059746                264 bp    mRNA    linear   EST 23-AUG-2000
DEFINITION  LE3a11.yg DNC15 Homo sapiens cDNA similar to Y BOX BINDING
            PROTEIN-1, mRNA sequence.
ACCESSION   AW059746
VERSION     AW059746.1  GI:6652068
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 264)
  AUTHORS   Brenner,S., Williams,S.R., Vermass,E.H., Storck,T., Moon,K.,
            McCollum,C., Mao,J.I., Kirchner,J.J., Eletr,S., DuBridge,R.B.,
            Burcham,T. and Albrecht,G.
  TITLE     In vitro cloning of complex mixtures of DNA on microbeads: Physical
            separation of differentially expressed cDNAs
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (4), 1665-1670 (2000)
   PUBMED   10677516
COMMENT     Contact: Burcham TS
            LYNX Therapeutics, Inc.
            25861 Industrial Blvd., Hayward, CA 94545, USA
            Tel: 510 670 9338
            Fax: 510 670 9302
            Email: timb@lynxgen.com
            Sequence obtained from LYNX Therapeutics Megasort technology.
            Collected from the down-regulated gate.
            High quality sequence stop: 264.
FEATURES             Location/Qualifiers
     source          1. .264
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /cell_type="monocyctic leukemia"
                     /cell_line="THP-1 (TIB-202)"
                     /clone_lib="DNC15"
                     /note="Vector: pCR2.1; Cloning of PCR products from
                     micro-beads carrying 3' end of down-regulated cDNA. THP-1
                     cells non-induced (treated with DMSO only). "
ORIGIN

  Query Match             100.0%;  Score 22;  DB 1;  Length 264;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        140 CTGCACAGGAGGGTTGGAATAC 161


RESULT 15
BG990119/c
LOCUS       BG990119                270 bp    mRNA    linear   EST 13-JUN-2001
DEFINITION  MR3-HT1104-230101-002-e01 HT1104 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BG990119
VERSION     BG990119.1  GI:14394189
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 270)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR3&t2=MR3-HT1104-
            230101-002-e01&t3=2001-01-23&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 270.
FEATURES             Location/Qualifiers
     source          1. .270
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /dev_stage="Adult"
                     /clone_lib="HT1104"
                     /note="Organ: head_neck; Vector: puc18; Site_1: SmaI;
                     Site_2: SmaI; A mini-library was made by cloning products
                     derived from  ORESTES PCR (U.S. Letters Patent application
                     No. 196,716 -  Ludwig Institute for Cancer Research)
                     profiles into  the pUC 18 vector. Reverse transcription of
                     tissue mRNA and cDNA amplification were performed under
                     low stringency conditions."
ORIGIN

  Query Match             100.0%;  Score 22;  DB 2;  Length 270;
  Best Local Similarity   100.0%;  Pred. No. 11;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db        197 CTGCACAGGAGGGTTGGAATAC 176


Search completed: January  8, 2006, 18:38:36
Job time : 2522 secs
```

GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

Run on: January 8, 2006, 16:24:54 ; Search time 306 Seconds
(without alignments)
479.161 Million cell updates/sec

Title: US-10-028-415-20
Perfect score: 22
Sequence: 1 ctgcacaggagggttggaatac 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	22	100.0	22	6	ABK88718	Abk88718 Human YB-
	2	22	100.0	152	10	ACD97151	Acd97151 Human col
c	3	22	100.0	250	6	AAI68908	Aai68908 Activated
	4	22	100.0	314	12	ACH84678	Ach84678 Human gen
c	5	22	100.0	364	10	ACD92102	Acd92102 Human col
	6	22	100.0	400	2	AAV87523	Aav87523 EST clone
	7	22	100.0	451	6	ABQ60291	Abq60291 Human col
c	8	22	100.0	478	9	ACH49466	Ach49466 Human leu
c	9	22	100.0	523	14	ACL62195	Acl62195 Human col
c	10	22	100.0	545	13	ADU12962	Adu12962 Solid tum
c	11	22	100.0	557	13	ACN38892	Acn38892 Tumour-as
c	12	22	100.0	560	6	ABQ57856	Abq57856 Human col
	13	22	100.0	595	12	ACH70978	Ach70978 Human gen
c	14	22	100.0	781	13	ACN39715	Acn39715 Tumour-as
	15	22	100.0	822	6	AAS62159	Aas62159 Porcine m
c	16	22	100.0	954	6	ABK50501	Abk50501 DNA encod
c	17	22	100.0	1353	13	ACN39713	Acn39713 Tumour-as
c	18	22	100.0	1363	12	ADQ84349	Adq84349 Human tum
c	19	22	100.0	1363	13	ACN37358	Acn37358 Tumour-as
c	20	22	100.0	1473	12	ADQ85063	Adq85063 Human tum
c	21	22	100.0	1473	13	ADQ87408	Adq87408 Human tum
c	22	22	100.0	1473	13	ACN38891	Acn38891 Tumour-as
c	23	22	100.0	1474	6	ABK84103	Abk84103 Human cDN
c	24	22	100.0	1474	10	ADG89384	Adg89384 Cancer de
c	25	22	100.0	1474	13	ADR25196	Adr25196 Breast ca
c	26	22	100.0	1474	13	ADP54732	Adp54732 Human PRO
c	27	22	100.0	1474	14	ADX97719	Adx97719 Human nuc
c	28	22	100.0	1474	14	ADX05879	Adx05879 Cyclin-de
c	29	22	100.0	1474	14	ADY15979	Ady15979 DNA encod
c	30	22	100.0	1481	4	AAC81319	Aac81319 Human Y-b
c	31	22	100.0	1481	6	ABS54998	Abs54998 Human YB-
c	32	22	100.0	1481	6	ABN95602	Abn95602 Gene #210
c	33	22	100.0	1481	10	ABZ99616	Abz99616 Human Y-b
c	34	22	100.0	1516	8	AAD55839	Aad55839 Human nuc
c	35	22	100.0	1520	10	ADB47382	Adb47382 Human cDN
c	36	22	100.0	1548	2	AAZ77507	Aaz77507 Human ova
c	37	22	100.0	1554	14	ADZ49037	Adz49037 Insulin s
c	38	22	100.0	1978	8	AAL51561	Aal51561 Human nuc
c	39	22	100.0	3073	2	AAQ14635	Aaq14635 Clone ass
c	40	22	100.0	3073	2	AAT34371	Aat34371 Plasmid p
c	41	22	100.0	3073	2	AAZ32246	Aaz32246 Human gli
c	42	22	100.0	3073	3	AAA88181	Aaa88181 pATG29 hu
c	43	20.4	92.7	448	8	ABX45462	Abx45462 Bovine ES
c	44	20.4	92.7	1404	13	ACN38702	Acn38702 Tumour-as
c	45	20.4	92.7	1522	13	ACN40311	Acn40311 Tumour-as

ALIGNMENTS

RESULT 1
ABK88718
ID ABK88718 standard; DNA; 22 BP.
XX
AC ABK88718;
XX
DT 07-OCT-2002 (first entry)
XX
DE Human YB-1 anti-sense strand, phosphorothioate oligonucleotide #2.
XX
KW Human; apoptotic cell death; proteinaceous transcription factor;
KW regulation of gene transcription; apoptosis; p53; CD95; TRA;
KW transcriptional regulator of apoptosis; Y-box family; YB-1; cancer;
KW tumour cell; embryonic cell; nervous system; intracellular pathogen;
KW DNA-damaging agent; retroviral infection; neurodegenerative disorder;
KW immune system dysfunction; anti-tumour; cytostatic; phosphorothioate; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT modified_base 1. .22
FT /*tag= a
FT /mod_base= OTHER
FT /note= "Phosphorothioate internucleotide linkages"
XX
PN WO200244363-A1.
XX
PD 06-JUN-2002.
XX
PF 28-NOV-2001; 2001WO-NZ000287.
XX
PR 28-NOV-2000; 2000US-00724809.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI Lasham A, Watson JD;
XX
DR WPI; 2002-557540/59.
XX
PT Modulating p53-mediated apoptotic cell death in a population of cells, by
PT modulating the amount of a transcriptional regulator of apoptosis
PT available to bind to a target polynucleotide in the cells.

XX
PS Example 2; Page 56; 62pp; English.
XX
CC The present invention relates to methods for modulating apoptotic cell
CC death using proteinaceous transcription factors that regulate the
CC transcription of genes encoding proteins involved in apoptosis (e.g. CD95
CC and p53). The methods involve modulating the amount of a transcriptional
CC regulator of apoptosis (TRA) available to bind to a target polynucleotide
CC in the cells, where TRA is a member of the Y-box nucleic acid binding
CC family of polypeptides (e.g. YB-1). The methods of the invention are
CC useful for modulating apoptotic cell death in a population of cells,
CC where the cells are selected from tumour cells, cells of the immune
CC system, embryonic cells, cells of the nervous system, or cells infected
CC with intracellular pathogens. The methods are also useful for increasing
CC the sensitivity of tumour cells to a DNA-damaging agent, and for
CC increasing sensitivity to apoptosis in a population of cells harbouring
CC intracellular pathogens. The methods are useful for screening an
CC apoptosis modulatory agent that modulates the binding of TRA. The methods
CC for regulating apoptosis can be used therapeutically and prophylactically
CC for various disorders such as cancer, viral and retroviral infections,
CC neurodegenerative disorders, and immune system dysfunction. The present
CC sequence represents a phosphorothioate oligonucleotide to the anti-sense
CC strand of human YB-1
XX
SQ Sequence 22 BP; 6 A; 4 C; 8 G; 4 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 6; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCACAGGAGGGTTGGAATAC 22
||||||||||||||||||||||
Db 1 CTGCACAGGAGGGTTGGAATAC 22

RESULT 2
ACD97151
ID ACD97151 standard; cDNA; 152 BP.
XX
AC ACD97151;
XX
DT 23-SEP-2003 (first entry)
XX
DE Human colon cancer cell expressed cDNA #5563.
XX
KW Open reading frame detection; genome sequencing; colon cancer;
KW breast cancer; population genome analysis; genetic shift; cancer;
KW antibiotic resistance; antibiotic non-tolerance; congenital disease;
KW agriculture; food crop genome; resistance gene; retrovirus;
KW influenza virus; eukaryotic pathogen detection; trypanosome; Plasmodium;
KW gene; ss.
XX
OS Homo sapiens.
XX
PN US2002155438-A1.
XX
PD 24-OCT-2002.
XX
PF 27-SEP-1999; 99US-00406117.
XX
PR 20-NOV-1998; 98US-00196716.
XX
PA (SIMP/) SIMPSON A J G.
PA (NETO/) NETO E D.
PA (BREN/) BRENTANI R R.
XX
PI Simpson AJG, Neto ED, Brentani RR;
XX
DR WPI; 2003-182626/18.
XX
PT Determining open reading frames of genome of an organism e.g. a human
PT suffering from cancer involves use of single oligonucleotide primer at
PT low stringency for preparing single-stranded cDNA from mRNA of
PT individual.
XX
PS Example 9; Page 794; 959pp; English.
XX
CC The invention describes a method of determining open reading frames in
CC the genome of organism, comprising contacting mRNA from cell of organism
CC with a single oligonucleotide primer (I) at low stringency, preparing
CC single-stranded cDNA by reverse transcribing mRNA with (I), amplifying
CC cDNA , sequencing the product, and repeating the contacting, preparing
CC and amplifying steps with different primers and sequencing resulting
CC nucleic acids. The method is useful for: determining that a known
CC nucleotide sequence from a genome of an organism corresponds to a
CC nucleotide sequence of an open reading frame; for preparing a contig,
CC nucleic acid molecule from a genome of an organism; and for sequencing
CC all or part of a genome of an organism. mRNA is obtained from mammalian
CC or human cell which is associated with a pathological condition e.g. a
CC colon cancer or breast cancer cell. The method is useful for analyses of
CC populations of subjects and can be used to carry out genetic analyses of
CC large or small populations. further, it can be used to study living
CC systems to determine if, e.g. there have been genetic shifts which render
CC an individual or population more or less likely to be afflicted with
CC diseases such as cancer, to determine antibiotic resistance or non-
CC tolerance, and so forth. The method can also be used in the study of
CC congenital diseases, and the risk of affliction to a foetus, as well as
CC the study of whether the conditions are likely to be passed to offspring
CC through ova or sperm. The analyses for pathological conditions can be
CC carried out in all animals, plants, birds, fish, etc. Using this method,
CC in the area of agriculture, for example the genomes of food crops can be
CC studied to determine if resistance genes are present, defects in plant
CC genomes can also be studied in this way. Similarly, the method permits
CC determination of the pathogens which integrate into the genome, such as
CC retroviruses and other integrating viruses such as influenza virus, have
CC undergone shifts or mutations, which may require different approaches to
CC therapy. This method is also applied to eukaryotic pathogens, such as
CC trypanosomes, different types of Plasmodium, etc. The method essentially
CC eliminates sequencing of non-coding portions. This sequence represents a
CC polynucleotide isolated from human colon cancer cell cDNA library
XX
SQ Sequence 152 BP; 27 A; 49 C; 34 G; 41 T; 0 U; 1 Other;

Query Match 100.0%; Score 22; DB 10; Length 152;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCACAGGAGGGTTGGAATAC 22
||||||||||||||||||||||
Db 87 CTGCACAGGAGGGTTGGAATAC 108

RESULT 3
AAI68908/c
ID AAI68908 standard; DNA; 250 BP.
XX
AC AAI68908;
XX
DT 29-JAN-2002 (first entry)
XX
DE Activated T-cell derived DNA fragment #44.
XX
KW Activated T-cell; immunosuppressive; immunostimulant; antiinflammatory;
KW cytostatic; gene therapy; vaccine; allergen; transplant rejection;
KW guest versus host disease; malignant disease; ds.
XX
OS Homo sapiens.
XX
PN DE10021834-A1.
XX
PD 15-NOV-2001.
XX
PF 06-MAY-2000; 2000DE-01021834.
XX
PR 06-MAY-2000; 2000DE-01021834.
XX

PA (LYNX-) LYNX THERAPEUTICS GMBH.
XX
PI Kramer MD, Winter H, Reinartz J;
XX
DR WPI; 2002-027320/04.
XX
PT New mRNA indicative of T cell activation and functional status, useful
PT for diagnosis and therapy e.g. of autoimmunity or transplant rejection.
XX
PS Claim 1; Page 20; 94pp; German.
XX
CC This sequence represents a novel messenger RNA, (mRNA), (I), for use as
CC indicator of the activation and functional status of T cells, that have
CC increased or reduced expression, and are present at higher or lower
CC concentration, in activated T cells, relative to normal or resting cells,
CC where (I) hybridizes to any of 334 sequences, reproduced, or their
CC derivatives, complements or fragments. The products of the invention have
CC immunosuppressive, immunostimulant, antiinflammatory and cytostatic
CC activity and can be used for gene therapy. The polynucleotides of the
CC invention are used: (i) as reagent for detecting activation/functional
CC status of T cells, for diagnosis, therapy, modulation or control of the
CC status, in cases of (auto)immunity (against microorganisms, vaccines or
CC allergens); transplant rejection; immunologically-related inflammation;
CC immunosuppression; immune deficiency; guest versus host disease, and
CC malignant diseases of the immune system; (ii) for identifying agents,
CC potential pharmaceuticals, that bind to (II) or derived polypeptides
CC (III); (iii) to prepare kits for measuring gene expression profiles in
CC isolated immune, especially T, cells; (iv) to raise antibodies (Ab)
CC directed against (III); and (v) to prepare binding molecules (IV)
CC specific for (II). Ab and (IV) are also useful for detecting and
CC modulating the activation and functional status of T cells. AAI68865-
CC AAI69198 represent the activated T-cell derived polynucleotide fragments
CC described in the method of the invention
XX
SQ Sequence 250 BP; 65 A; 71 C; 77 G; 37 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 6; Length 250;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCACAGGAGGGTTGGAATAC 22
||||||||||||||||||||||
Db 125 CTGCACAGGAGGGTTGGAATAC 104

RESULT 4
ACH84678
ID ACH84678 standard; DNA; 314 BP.
XX
AC ACH84678;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon probe #17873.
XX
KW Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
PN US2003194704-A1.
XX
PD 16-OCT-2003.
XX
PF 03-APR-2002; 2002US-00029386.
XX
PR 03-APR-2002; 2002US-00029386.
XX
PA (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX
PI Penn SG, Rank DR, Hanzel DK;
XX
DR WPI; 2004-119264/12.
XX
PT New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
PS Claim 1; SEQ ID NO 17873; 80pp; English.
XX
CC The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above- mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ Sequence 314 BP; 58 A; 83 C; 83 G; 90 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 12; Length 314;
Best Local Similarity 100.0%; Pred. No. 0.73;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCACAGGAGGGTTGGAATAC 22
||||||||||||||||||||||
Db 43 CTGCACAGGAGGGTTGGAATAC 64

RESULT 5
ACD92102/c
ID ACD92102 standard; cDNA; 364 BP.
XX
AC ACD92102;
XX
DT 23-SEP-2003 (first entry)
XX
DE Human colon cancer cell expressed cDNA #514.
XX
KW Open reading frame detection; genome sequencing; colon cancer;
KW breast cancer; population genome analysis; genetic shift; cancer;
KW antibiotic resistance; antibiotic non-tolerance; congenital disease;
KW agriculture; food crop genome; resistance gene; retrovirus;
KW influenza virus; eukaryotic pathogen detection; trypanosome; Plasmodium;
KW gene; ss.

XX
OS Homo sapiens.
XX
PN US2002155438-A1.
XX
PD 24-OCT-2002.
XX
PF 27-SEP-1999; 99US-00406117.
XX
PR 20-NOV-1998; 98US-00196716.
XX
PA (SIMP/) SIMPSON A J G.
PA (NETO/) NETO E D.
PA (BREN/) BRENTANI R R.
XX
PI Simpson AJG, Neto ED, Brentani RR;
XX
DR WPI; 2003-182626/18.
XX
PT Determining open reading frames of genome of an organism e.g. a human
PT suffering from cancer involves use of single oligonucleotide primer at
PT low stringency for preparing single-stranded cDNA from mRNA of
PT individual.
XX
PS Example 9; Page 76; 959pp; English.
XX
CC The invention describes a method of determining open reading frames in
CC the genome of organism, comprising contacting mRNA from cell of organism
CC with a single oligonucleotide primer (I) at low stringency, preparing
CC single-stranded cDNA by reverse transcribing mRNA with (I), amplifying
CC cDNA , sequencing the product, and repeating the contacting, preparing
CC and amplifying steps with different primers and sequencing resulting
CC nucleic acids. The method is useful for: determining that a known
CC nucleotide sequence from a genome of an organism corresponds to a
CC nucleotide sequence of an open reading frame; for preparing a contig,
CC nucleic acid molecule from a genome of an organism; and for sequencing
CC all or part of a genome of an organism. mRNA is obtained from mammalian
CC or human cell which is associated with a pathological condition e.g. a
CC colon cancer or breast cancer cell. The method is useful for analyses of
CC populations of subjects and can be used to carry out genetic analyses of
CC large or small populations. further, it can be used to study living
CC systems to determine if, e.g. there have been genetic shifts which render
CC an individual or population more or less likely to be afflicted with
CC diseases such as cancer, to determine antibiotic resistance or non-
CC tolerance, and so forth. The method can also be used in the study of
CC congenital diseases, and the risk of affliction to a foetus, as well as
CC the study of whether the conditions are likely to be passed to offspring
CC through ova or sperm. The analyses for pathological conditions can be
CC carried out in all animals, plants, birds, fish, etc. Using this method,
CC in the area of agriculture, for example the genomes of food crops can be
CC studied to determine if resistance genes are present, defects in plant
CC genomes can also be studied in this way. Similarly, the method permits
CC determination of the pathogens which integrate into the genome, such as
CC retroviruses and other integrating viruses such as influenza virus, have
CC undergone shifts or mutations, which may require different approaches to
CC therapy. This method is also applied to eukaryotic pathogens, such as
CC trypanosomes, different types of Plasmodium, etc. The method essentially
CC eliminates sequencing of non-coding portions. This sequence represents a
CC polynucleotide isolated from human colon cancer cell cDNA library
XX
SQ Sequence 364 BP; 113 A; 105 C; 99 G; 47 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 10; Length 364;
Best Local Similarity 100.0%; Pred. No. 0.74;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCACAGGAGGGTTGGAATAC 22
||||||||||||||||||||||
Db 105 CTGCACAGGAGGGTTGGAATAC 84

RESULT 6
AAV87523
ID AAV87523 standard; cDNA; 400 BP.
XX
AC AAV87523;
XX
DT 12-FEB-1999 (first entry)
XX
DE EST clone BK8.
XX
KW Expressed sequence tag; secreted protein; haematopoiesis regulator;
KW tissue growth; activin; inhibin; tumour invasion suppressor; EST; human;
KW chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
KW receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
XX
OS Homo sapiens.
XX
PN WO9845437-A2.
XX
PD 15-OCT-1998.
XX
PF 10-APR-1998; 98WO-US006956.
XX
PR 10-APR-1997; 97US-00837312.
XX
PA (GEMY) GENETICS INST INC.
XX
PI Jacobs K, Mccoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;
PI Spaulding V, Agostino MJ;
XX
DR WPI; 1999-070078/06.
XX
PT New polynucleotides encoding human secreted proteins - derived from e.g.
PT human blood, kidney, foetal lung, placenta, testes, brain, ovary,
PT pituitary, retina and colon cDNA libraries.
XX
PS Claim 1; Page 92-93; 641pp; English.
XX
CC The present sequence represents an expressed sequence tag (EST), and is a
CC polynucleotide of the invention. The polynucleotides of the invention are
CC all secreted EST sequences isolated from a variety of human tissue
CC sources. The EST sequences and proteins encoded by them are predicted to
CC have useful biological activities which would make them suitable for
CC treating, preventing or ameliorating medical conditions in humans and
CC animals, although no supporting data is given. Suggested activities
CC include nutritional activity, immune stimulating or suppressing activity,
CC haematopoiesis regulating activity, tissue growth activity,
CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC activity, cadherin/tumour invasion suppressor activity, tumour inhibition
CC activity. The EST sequences are also stated to be useful for gene therapy
XX
SQ Sequence 400 BP; 60 A; 112 C; 107 G; 121 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 2; Length 400;
Best Local Similarity 100.0%; Pred. No. 0.75;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCACAGGAGGGTTGGAATAC 22
||||||||||||||||||||||
Db 200 CTGCACAGGAGGGTTGGAATAC 221

RESULT 7
ABQ60291
ID ABQ60291 standard; cDNA; 451 BP.
XX
AC ABQ60291;
XX
DT 02-AUG-2002 (first entry)
XX
DE Human colon cancer related nucleotide sequence SEQ ID NO:3986.
XX
KW Human; colon cancer; cancer; tissue profiling; forensic; mapping;
KW genetic analysis; diagnostic; antisense therapy; gene; ss.

XX
OS Homo sapiens.
XX
PN WO200229086-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US030732.
XX
PR 02-OCT-2000; 2000US-0237271P.
XX
PA (FARB) BAYER CORP.
XX
PI Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
PI Thiaglingam A, Lewis ME;
XX
DR WPI; 2002-426115/45.
XX
PT New isolated nucleic acid that is differentially expressed in cancer
PT tissues useful for determining the presence of colon cancer in a cell or
PT tissue type, and in antisense therapy.
XX
PS Claim 1; Fig 1; 796pp; English.
XX
CC ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
CC expressed in cancer tissues. ABB78993 to ABB79004 represent proteins
CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
CC used in antisense therapy. An antibody immunoreactive with a polypeptide
CC encoded by (I) is useful for detecting cancer in a patient sample, and
CC for detecting the presence or absence of a polynucleotide encoded by a
CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived
CC from (I) can be used for determining the presence of a nucleic acid which
CC hybridises to (I), and for determining the phenotype of cells in a sample
CC of cells from a patient. (I) is useful for determining the presence of
CC colon cancer in a cell or tissue type, for determining the presence or
CC state of other type of cancer, in antisense therapy, to generate
CC macroarrays on a solid surface, to identify a chromosome on which the
CC corresponding gene resides, and in tissue profiling, forensics, genetic
CC analysis, mapping and diagnostic applications. (I) can be used to raise
CC antibodies, and to screen for peptide analogues and antagonists
XX
SQ Sequence 451 BP; 68 A; 129 C; 115 G; 137 T; 0 U; 2 Other;

Query Match 100.0%; Score 22; DB 6; Length 451;
Best Local Similarity 100.0%; Pred. No. 0.76;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCACAGGAGGGTTGGAATAC 22
||||||||||||||||||||||
Db 200 CTGCACAGGAGGGTTGGAATAC 221

RESULT 8
ACH49466/c
ID ACH49466 standard; cDNA; 478 BP.
XX
AC ACH49466;
XX
DT 13-OCT-2003 (first entry)
XX
DE Human leukocyte cDNA #1060.
XX
KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW genome mapping; biodiversity; genetic disorder.
XX
OS Homo sapiens.
XX
PN US2003073623-A1.
XX
PD 17-APR-2003.
XX
PF 30-JUL-2001; 2001US-00918995.
XX
PR 30-JUL-2001; 2001US-00918995.
XX
PA (DRMA/) DRMANAC R T.
PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
XX
PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX
DR WPI; 2003-615964/58.
XX
PT New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.
XX
PS Claim 1; SEQ ID NO 36678; 44pp; English.
XX
CC The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversities, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623
XX
SQ Sequence 478 BP; 139 A; 129 C; 128 G; 66 T; 0 U; 16 Other;

Query Match 100.0%; Score 22; DB 9; Length 478;
Best Local Similarity 100.0%; Pred. No. 0.76;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCACAGGAGGGTTGGAATAC 22
||||||||||||||||||||||
Db 101 CTGCACAGGAGGGTTGGAATAC 80

RESULT 9
ACL62195/c
ID ACL62195 standard; cDNA; 523 BP.
XX
AC ACL62195;
XX
DT 24-MAR-2005 (first entry)
XX
DE Human colon cancer differentially expressed polynucleotide, SEQ ID:8330.
XX
KW Differential expression; diagnosis; therapy; drug screening; cancer;
KW neoplasm; colon tumor; breast tumor; pancreas tumor; cytostatic; vaccine;
KW ss.
XX
OS Homo sapiens.
XX
PN WO2005000087-A2.
XX
PD 06-JAN-2005.
XX
PF 13-MAY-2004; 2004WO-US015421.
XX
PR 03-JUN-2003; 2003US-0475872P.
XX
PA (CHIR) CHIRON CORP.

XX
PI Randazzo F, Moler E, Escobedo J, Garcia PD;
XX
DR WPI; 2005-075421/08.
XX
PT New isolated polynucleotides, which are differentially expressed in colon
PT cancer cell, useful for treating cancer, e.g. colon cancer, breast
PT cancer, or pancreatic cancer.
XX
PS Claim 1; SEQ ID NO 8330; 97pp; English.
XX
CC The invention relates to 9672 polynucleotides (ACL53866-ACL63537) which
CC are differentially expressed in colon cancer cells. The invention also
CC relates to vectors and host cells comprising a differentially expressed
CC polynucleotide of the invention; a method for detecting a cancerous cell
CC by detection of a gene product of the polynucleotides; a method for
CC inhibiting a cancerous phenotype of a cell by inhibiting a gene product
CC of the polynucleotides; a method of treating an individual with cancer by
CC administration of a modulator of a gene product of the polynucleotides;
CC and an isolated antibody that specifically binds to a polypeptide encoded
CC by one of the 9672 polynucleotides. The polynucleotides, polypeptides,
CC antibodies, and methods are useful for the detection of cancerous cells;
CC for the diagnosis, prognosis and management of cancer; for the
CC identification of agents that modulate the phenotype of cancerous cells;
CC for the identification of therapeutic targets for cancer chemotherapy;
CC and for the treatment of cancer, especially colon cancer and metastasized
CC colon cancer, but also breast or pancreatic cancer. The polynucleotides
CC are also useful as a source of probes or primers for use in diagnostic
CC methods. The differentially expressed polynucleotides or their encoded
CC proteins can additionally be used as vaccines to modulate primary immune
CC responses for the prevention or treatment of cancer. The present sequence
CC represents a specifically claimed polynucleotide which is differentially
CC expressed in colon cancer. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 523 BP; 158 A; 112 C; 153 G; 100 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 14; Length 523;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCACAGGAGGGTTGGAATAC 22
||||||||||||||||||||||
Db 487 CTGCACAGGAGGGTTGGAATAC 466

RESULT 10
ADU12962/c
ID ADU12962 standard; DNA; 545 BP.
XX
AC ADU12962;
XX
DT 27-JAN-2005 (first entry)
XX
DE Solid tumour prognosis gene seqid 3401.
XX
KW cytostatic; gene therapy; expression profile; solid tumour;
KW peripheral blood mononuclear cell; PBMC; prognosis; ds.
XX
OS Unidentified.
XX
PN WO2004097052-A2.
XX
PD 11-NOV-2004.
XX
PF 29-APR-2004; 2004WO-US013587.
XX
PR 29-APR-2003; 2003US-0466067P.
PR 23-JAN-2004; 2004US-0538246P.
XX
PA (AMHP) WYETH.
PA (STRA/) STRAHS A.
XX
PI Strahs A, Trepicchio WL, Burczynski ME, Twine NC, Slonim DK;
PI Immerman F, Dorner AJ;
XX
DR WPI; 2004-804779/79.
XX
PT A method, useful for prognosing and treating solid tumor, comprises
PT comparing an expression profile of a gene expressed in peripheral blood
PT mononuclear cells to a reference expression profile of a gene.
XX
PS Disclosure; Page; 111pp; English.
XX
CC The invention describes a method comprising comparing an expression
CC profile of at least one gene in a peripheral blood sample of a patient to
CC at least one reference expression profile of the at least one gene, where
CC the patient has a solid tumour, and each of the gene is differentially
CC expressed in peripheral blood mononuclear cells (PBMCs) of a first class
CC of patients as compared to PBMCs of a second class of patients, where
CC both the first and second classes of patients have the solid tumour, and
CC each of the first and second classes is a subcluster formed by an
CC unsupervised clustering analysis of gene expression profiles in PBMCs of
CC a population of patients who have the solid tumour, and where the
CC majority of the first class of patients has a first clinical outcome, and
CC the majority of the second class of patients has a second clinical
CC outcome. Also described are: a system comprising (i) a memory or a
CC storage medium including data that represent an expression profile of at
CC least one gene in a peripheral blood sample of a patient who has a solid
CC tumour, (ii) at least another storage medium including data that
CC represent at least one reference expression profile of the gene, (iii) a
CC program capable of comparing the expression profile to the reference
CC expression profile, and (iv) a processor capable of executing the
CC program, where expression levels of the gene in peripheral blood
CC mononuclear cells of patients who have the solid tumour correlate with
CC clinical outcomes of the patients; and a nucleic acid or protein array
CC comprising concentrated probes for solid tumour prognosis genes, where
CC each of the solid tumour prognosis genes is differentially expressed in
CC PBMCs of a first class of patients as compared to PBMCs of a second class
CC of patients, where both the first and second classes of patients have a
CC solid tumour, and where the first class of patients has a first clinical
CC outcome, and the second class of patients has a second clinical outcome.
CC The method, system, and array are useful for prognosing and treating
CC solid tumours. This sequence represents a solid tumour prognosis gene of
CC the invention. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 545 BP; 167 A; 120 C; 134 G; 90 T; 0 U; 34 Other;

Query Match 100.0%; Score 22; DB 13; Length 545;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCACAGGAGGGTTGGAATAC 22
||||||||||||||||||||||
Db 36 CTGCACAGGAGGGTTGGAATAC 15

RESULT 11
ACN38892/c
ID ACN38892 standard; cDNA; 557 BP.
XX
AC ACN38892;
XX
DT 18-NOV-2004 (first entry)
XX
DE Tumour-associated antigenic target (TAT) cDNA DNA325180, SEQ ID NO:2688.
XX
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;

KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2004030615-A2.
XX
PD 15-APR-2004.
XX
PF 29-SEP-2003; 2003WO-US028547.
XX
PR 02-OCT-2002; 2002US-0414971P.
XX
PA (GETH) GENENTECH INC.
XX
PI Wu TD, Zhang Z, Zhou Y;
XX
DR WPI; 2004-347921/32.
XX
PT New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.
XX
PS Claim 1; SEQ ID NO 2688; 7273pp; English.
XX
CC The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with
CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT nucleic acid of the invention
XX
SQ Sequence 557 BP; 182 A; 133 C; 131 G; 111 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 13; Length 557;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCACAGGAGGGTTGGAATAC 22
||||||||||||||||||||||
Db 51 CTGCACAGGAGGGTTGGAATAC 30

RESULT 12
ABQ57856/c
ID ABQ57856 standard; cDNA; 560 BP.
XX
AC ABQ57856;
XX
DT 02-AUG-2002 (first entry)
XX
DE Human colon cancer related nucleotide sequence SEQ ID NO:1551.
XX
KW Human; colon cancer; cancer; tissue profiling; forensic; mapping;
KW genetic analysis; diagnostic; antisense therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200229086-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US030732.
XX
PR 02-OCT-2000; 2000US-0237271P.
XX
PA (FARB) BAYER CORP.
XX
PI Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
PI Thiaglingam A, Lewis ME;
XX
DR WPI; 2002-426115/45.
XX
PT New isolated nucleic acid that is differentially expressed in cancer
PT tissues useful for determining the presence of colon cancer in a cell or
PT tissue type, and in antisense therapy.
XX
PS Claim 1; Fig 1; 796pp; English.
XX
CC ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
CC expressed in cancer tissues. ABB78993 to ABB79004 represent proteins
CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
CC used in antisense therapy. An antibody immunoreactive with a polypeptide
CC encoded by (I) is useful for detecting cancer in a patient sample, and
CC for detecting the presence or absence of a polynucleotide encoded by a
CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived
CC from (I) can be used for determining the presence of a nucleic acid which
CC hybridises to (I), and for determining the phenotype of cells in a sample
CC of cells from a patient. (I) is useful for determining the presence of
CC colon cancer in a cell or tissue type, for determining the presence or
CC state of other type of cancer, in antisense therapy, to generate
CC macroarrays on a solid surface, to identify a chromosome on which the
CC corresponding gene resides, and in tissue profiling, forensics, genetic
CC analysis, mapping and diagnostic applications. (I) can be used to raise
CC antibodies, and to screen for peptide analogues and antagonists
XX
SQ Sequence 560 BP; 166 A; 128 C; 164 G; 97 T; 0 U; 5 Other;

Query Match 100.0%; Score 22; DB 6; Length 560;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCACAGGAGGGTTGGAATAC 22
||||||||||||||||||||||
Db 348 CTGCACAGGAGGGTTGGAATAC 327

RESULT 13
ACH70978
ID ACH70978 standard; DNA; 595 BP.
XX
AC ACH70978;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon probe #4173.
XX
KW Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
PN US2003194704-A1.
XX
PD 16-OCT-2003.
XX
PF 03-APR-2002; 2002US-00029386.
XX
PR 03-APR-2002; 2002US-00029386.
XX
PA (PENN/) PENN S G.

PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX
PI Penn SG, Rank DR, Hanzel DK;
XX
DR WPI; 2004-119264/12.
XX
PT New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
PS Claim 15; SEQ ID NO 4173; 80pp; English.
XX
CC The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above- mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ Sequence 595 BP; 157 A; 153 C; 121 G; 164 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 12; Length 595;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCACAGGAGGGTTGGAATAC 22
||||||||||||||||||||||
Db 299 CTGCACAGGAGGGTTGGAATAC 320

RESULT 14
ACN39715/c
ID ACN39715 standard; cDNA; 781 BP.
XX
AC ACN39715;
XX
DT 18-NOV-2004 (first entry)
XX
DE Tumour-associated antigenic target (TAT) cDNA DNA325908, SEQ ID NO:4034.
XX
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2004030615-A2.
XX
PD 15-APR-2004.
XX
PF 29-SEP-2003; 2003WO-US028547.
XX
PR 02-OCT-2002; 2002US-0414971P.
XX
PA (GETH) GENENTECH INC.
XX
PI Wu TD, Zhang Z, Zhou Y;
XX
DR WPI; 2004-347921/32.
XX
PT New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.
XX
PS Claim 1; SEQ ID NO 4034; 7273pp; English.
XX
CC The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with
CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT nucleic acid of the invention
XX
SQ Sequence 781 BP; 218 A; 208 C; 233 G; 122 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 13; Length 781;
Best Local Similarity 100.0%; Pred. No. 0.81;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGCACAGGAGGGTTGGAATAC 22
||||||||||||||||||||||
Db 663 CTGCACAGGAGGGTTGGAATAC 642

RESULT 15
AAS62159
ID AAS62159 standard; cDNA; 822 BP.
XX
AC AAS62159;
XX
DT 29-JAN-2002 (first entry)
XX
DE Porcine muscular steatosis-modulating factor #285.
XX
KW Pig; muscular steatosis-modulating factor; ss; metabolic; muscular; MSMF;
KW food supplement; obesity; hyperlipidaemia; atherosclerosis;

KW wound healing; tumour; amyotrophic lateral sclerosis; ALS.
XX
OS Sus scrofa.
XX
PN WO200179287-A2.
XX
PD 25-OCT-2001.
XX
PF 12-APR-2001; 2001WO-CA000509.
XX
PR 17-APR-2000; 2000US-0197936P.
XX
PA (MIAC) CANADA AGRIC & AGRI-FOOD CANADA.
XX
PI Palin M, Pomar C, Gariepy C;
XX
DR WPI; 2002-017600/02.
XX
PT Prognosis and diagnosis of muscular steatosis, useful e.g. for selecting
PT animals for breeding, by measuring levels of specific markers, also
PT treating or inducing steatosis.
XX
PS Claim 5; Page 182; 190pp; English.
XX
CC The invention relates to prognosis or diagnosis of muscular steatosis by
CC measuring the level of a muscular steatosis modulating factor (MSMF) in a
CC human or animal and comparing this with the level in a healthy control.
CC Any difference indicates presence of, or predisposition to, muscular
CC steatosis. The method is particularly used for diagnosis or prognosis of
CC muscular steatosis in mammals and birds, e.g. to select individuals as
CC founders in animal breeding. Also (ant)agonists of MSMF can be used to
CC treat, or induce (for increasing the fat content of food) muscular
CC steatosis, in humans and animals. The MSMF markers are also useful in the
CC study of diseases and conditions such as obesity, hyperlipidaemia,
CC atherosclerosis, wound healing, tumours and amyotrophic lateral sclerosis
CC (ALS). The present sequence is a MSMF of the invention
XX
SQ Sequence 822 BP; 155 A; 224 C; 177 G; 212 T; 0 U; 54 Other;

Query Match 100.0%; Score 22; DB 6; Length 822;
Best Local Similarity 100.0%; Pred. No. 0.81;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 CTGCACAGGAGGGTTGGAATAC 22
              ||||||||||||||||||||||
Db         71 CTGCACAGGAGGGTTGGAATAC 92
```

Search completed: January 8, 2006, 17:27:30
Job time : 309 secs